

STIC-Biotech/ChemLib

From: Helmer, Georgia
Sent: Wednesday, August 28, 2002 10:51 AM
To: STIC-Biotech/ChemLib
Subject: FW: seq search 09/810861

-----Original Message-----

From: Helmer, Georgia
Sent: Wednesday, August 28, 2002 10:33 AM
T : STIC-ILL
Subject: seq search 09/810861

Arntzen

Please do DNA sequence searches, and oligo searches, for SEQ ID NOs 1 thru 5 of this case.

Thanks in advance!

Georgia L. Helmer Ph.D.
Patent Examiner
Crystal Mall 1, 9D14
AU 1638
703-308-7023
Georgia.Helmer@USPTO.gov
mailbox 9e12

12/1/02

12/1/02

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6806 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 8/30
Date Completed: 9/3/02
Searcher Prep/Review: 10
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: 4
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:41:47 ; Search time 9324.54 Seconds
(without alignments)
3871.320 Million cell updates/sec

Title: US-09-810-861B-5

Perfect score: 1725
Sequence: 1 atgagggcccccagcagctctc.....aattgctcagcgccacctga 1725

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_de:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ats:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ats:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	1725	100.0	1725	6	AX275256	AX275256 Sequence
2	1722	99.8	2218	9	HUMACHE	M55040 Human acetyl
3	1722	99.8	2256	6	AR070205	AR070205 Sequence
4	1722	99.8	3016	6	AR070207	AR070207 Sequence
5	1722	99.8	3096	6	AR070206	AR070206 Sequence
6	1712	99.2	5767	6	AX275254	AX275254 Sequence
7	1712	99.2	5767	6	AX275255	AX275255 Sequence
8	1070	62.0	1355	9	HUMACHE01	L22559 Human acetyl
9	1070	62.0	4185	9	HUMACHEB	L42812 Homo sapien
10	1070	62.0	172358	9	AC011895	AC011895 Homo sapi
11	1070	62.0	194681	2	AC084057	AC084057 Homo sapi
12	781	45.3	34921	9	HSF002993	AF002993 Homo sapi
13	489	28.3	779	9	HUMACHE02	L22560 Human acetyl
14	170	9.9	510	9	HUMACHE03	L22561 Human acetyl
15	164	9.5	117962	9	AF312032	AF312032 Homo sapi
16	60	3.5	3958	4	AF053485	AF053485 Fells cat
17	53	3.1	1947	4	OCU05036	U05036 Oryctolagus
18	48	2.8	503	4	BTACHE2	AF061814 Bos tauru
19	47	2.7	984	4	BTACHE1	AF061813 Bos tauru
20	33	1.9	2089	6	AX306139	AX306139 Sequence
21	33	1.9	2089	10	MMACHE	X56518 Mouse mRNA
22	33	1.9	296820	10	AF312033	AF312033 Mus muscu
23	32	1.9	2066	10	S50879	S50879 acetylcholi
24	30	1.7	535	9	HUMACHEA	M76539 Human acetyl
25	28	1.6	18385	9	AC009084	AC009084 Homo sapi
26	28	1.6	206102	9	AC027131	AC027131 Homo sapi
27	27	1.6	1538	6	AX180336	AX180336 Sequence
28	25	1.4	2523	6	AX179306	AX179306 Sequence
29	25	1.4	3993	10	RNU41662	U41662 Rattus norv
30	25	1.4	65064	2	AC098752	AC098752 Rattus no
31	25	1.4	146756	2	AC002348	AC002348 Homo sapi
32	25	1.4	185010	2	AC022514	AC022514 Mus muscu
33	25	1.4	190358	2	AC016876	AC016876 Homo sapi
34	25	1.4	222871	2	AC108094	AC108094 Homo sapi
35	25	1.4	234182	10	AL603707	AL603707 Mouse DNA
36	24	1.4	840	10	MUSACHEA	M76540 Mouse acetyl
37	23	1.3	667	10	AF151982	AF151982 Rattus no
38	23	1.3	182844	2	AL671857	AL671857 Mus muscu
39	22	1.3	3945	10	AB043123	AB043123 Mus muscu
40	22	1.3	4011	10	AB043125	AB043125 Mus muscu
41	22	1.3	4014	10	AB043124	AB043124 Mus muscu
42	22	1.3	4173	10	AB005662	AB005662 Mus muscu
43	22	1.3	4545	10	AF262046	AF262046 Mus muscu
44	22	1.3	5443	10	AF178636	AF178636 Mus muscu
45	22	1.3	5532	10	AF178637	AF178637 Mus muscu

ALIGNMENTS

RESULT 1
LOCUS AX275256 1725 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 5 from Patent WO0171014.
ACCESSION AX275256
VERSION AX275256.1 GI:16547676
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (sites)
MOR.T., Soreq,H., Arntzen,C. and Mason,H.
Expression of recombinant human acetylcholinesterase in transgenic
plants
Patent: WO 0171014-A 5 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsafirir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source
1..1725
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic human acetylcholinesterase gene optimized"

	for expression in plants"			
BASE COUNT	273 a	553 c	571 g	328 t
ORIGIN				

Query Match	100.0%	Score 1725;	DB 6;	Length 1725;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1725; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	atgagagcccccagagatgtctctgtgcacagacccctcccttgctcccccacatccctctccctc	60
Db	1	ATGAGGCCCCCGAGAGTGTCTGTCTGCACAGCCCTTCCCTGGCTTCCCTCCACATCTTCTCTC	60
QY	61	ctccctctgactccctcgggtctgagagagctgaggggtctgagagccggagagatgacagagctgtc	120
Db	61	CTCCCTCTGCTCTCTGGGTGTGAGAGATGTGGGGGCTGTAGGGCCGGAGATGTCAAGACTGTCTG	120
QY	121	gtgacagctgctgagggagccggtctgcgggagcttcgccttgaaagaccccccgggggacctgtc	180
Db	121	GTGACAGGTCGTGGGGGGCCGGCTCGGGGGCAATTGCTCTGAAGAACCCCGGGGGCCCTGTCTC	180
QY	181	ctctgtctctctgaggacatccctcttgcgagagccacacatgagaccccgctgtctctgcga	240
Db	181	CTCTCTCTTCTCTGGGCAATCCCTTTGTGCGAGACACCCCTATGGGACCCCGCTCTTCTGTGCCA	240
QY	241	ccggagaccgaagagacctgtgtacaggggtgtgtagaagctaacacttccagagtgatctgc	300
Db	241	CCGGAGCCCAAGACAGCCTTGGTGTACGGGGTGTGTAGACCTTCAACCTTCCAGATGTCTGTC	300
QY	301	taccacatagtctgagacacccctataccacaggtcttgtagggcacccgagatctgtaaaccccaac	360
Db	301	TACCAATATGTGTGACACCCCTATACCAGGTTTGTAGGGGCACCGAGATGTGAACCCCAAC	360
QY	361	cgtagagctgagccggagagctgtcgtcttaactaaagctgtgagacacataccccgggtctaca	420
Db	361	CGTAGAGCTGAGCCGAGAGCTGCTGTACTTAAGGTGTGGACACCATACCCCGGGCTTACA	420
QY	421	tcccccacaccctgtccctcgtctctgatactatctaggggtgtgctctctacagttagggctctcc	480
Db	421	TCCCCACACCCCTGTCTCTGTCCTGTGATGTAGGGGGTGTGCTTTCATCAATGGGGCTCTCTCC	480
QY	481	cttgagagctgtaagatatggccgctctctctgtactcaaggccgagagagacgtgtcgtgtgctcatg	540
Db	481	CTTGAGACGTGTACGATGTGGCCGCTTCTTGTTGTACAGGCCAGAGAGGACTGTGCTGTCTCATG	540
QY	541	aacacacccgggtggagagaccttggtctctctctgtgacctgcggggagccgagagggccccgggc	600
Db	541	AACATACCGGGGGGAGAGCTTGTGGCTTCTGTGCCCCGTGCCGGGGAGCCGAGAGGGCCCCGGGC	600
QY	601	aaatgtggatctctctctcgtgatacagagctgtgcacctgtgagctgtgacaggaagacgtgtgcacc	660
Db	601	AATGTGGGTCTCTCGATGTACAGAGGCTTGCCCTGCAGTGGGTGCAGGAGAACTGTGCGACCC	660
QY	661	ctcgggggttgagaccgcagacatcagtgagcgtgtcttgaggagagacgctggagagccgtctcgtgtg	720
Db	661	CTCGGGGGGTGACCCGACATCATGATGACCTGTTGGGGAGAGACCGGGAGCGGCTCGTCTGCTG	720
QY	721	ggatctgacactcgtgtctcccgcccaagccggggcgtgtcttcaacagggcgtgtgtctgcagagac	780
Db	721	GGCATGTGACCTGTGCTGTCCCGCCACAGCCGGGGGCTGTCTTCCACAGGGCCGTGTGTGCAGAGAC	780
QY	781	ggatcccccacaaatgagaccctctggccacagctgagctatgtagaagagcccgctgcagagggcacag	840
Db	781	GGTACCCCCCAATGTGACCTTGGGCCACGGGTGTGGCATGTGGAGAGGCCCTGTCTCAAGGGCCACG	840
QY	841	caagcttgacccaactctgttgagcgtgtctctcagcagcgcaactgtgtggaatgacacagaagactgt	900
Db	841	CAGCTGGCCCACTTGTGGGCTGTCTCTCCAGCGGCACACTGTGTGGATGACACAGAGCTGTG	900
QY	901	gttagccctgactctggagacagacagcgcaagcttccgtgtgtgaacagaaatgacagctgtgtg	960
Db	901	GTATCCCTGTCTTGTGGACACAGACAGCGCCAGAGTCTCTGTGTAAACACAGAAATGTGACGTGTGTG	960

QY	961	cctaagaagaagcgtctccggtctctccgtctcgtctgctgtgtagaagtgaaagacttcccaagt	1020
Db	961	CCCTCAAGAAAGCGCTCTCCGCTTCTCTTCGTGCTGTGGTATGATGAGAGATTCCTCAGT	1020
QY	1021	gaacccccaaaggccctcatctcaacgvgggaagcttccacggtccctgacaggtgtctgtgtgt	1080
Db	1021	GACACCCCAAGAGCCCTCATCAACGGGGGAAGCTTCACAGGCTGTCAAGTCTGTGGGT	1080
QY	1081	gtgtgtgaagaagtatgaaggtctcgttaattctcgtgtttacgvggccccagagcttcaagaagac	1140
Db	1081	GTGTGTGAAGATGAGGGCTGTATTTTCGTGTTACGGGGCCCCACGAGTTCAGCAAAAGAC	1140
QY	1141	aacgagctctcatcacagccgggagccgaagcttccgtgcccgggtgtcggtgtgtgtgtcccaag	1200
Db	1141	AACGAGCTCTCATCAACGGGGCCGAGTTTCTGTGGCGGGGTGTGGGGTTCCTCCACAG	1200
QY	1201	gtaaagtaccttgcacgacggaaggtgtgtgtcctgtacgtatcaacaaagactgtctcatcccgag	1260
Db	1201	GTAAGTACCTTGGCAGCCGAGGCTGTGTGCTCTGCAATTACACAGATGGAGTCAATCCGAG	1260
QY	1261	gaacccgacacgcctgtgaaggagggcccttgaagcgaatgtgtgtgtgtgtgtgtgtgtgtgtgc	1320
Db	1261	GACCCGGCAGCGCTGAAGGAGGGCCGTGACGATGTGTGTGTGTGTGTGTGTGTGTGTGC	1320
QY	1321	cccggtgcccagctgtgctgtgtgtgcagtgtgtgtccaaaggtgtccgggtgtctacgctacgttc	1380
Db	1321	CCCGTGGCCACAGTGTGCTGTGGCCACAGCTGTGCCAAGGGTGTCCCGGTGTACGCTTACGTC	1380
QY	1381	tttgaacacggtcttctcacaagctctctctgtgcccctgtgaattgtgtgtgtgtgtgtgtgtgtac	1440
Db	1381	TTTGAACACCGTCTTCCACGCTCTCTCTGTGCCCCCTGTGGATGTGGGGGTGTGCCCAAGGCTTAC	1440
QY	1441	gagaatcgaattcatcttcttggatctccccccttgcacccctctcgaataactaaacgacgaagagag	1500
Db	1441	GAGATCGAGTTCAATCTTTGGGATCCCCCTGTGACCCCTCTCGAAACTACACGGCAGAGAGAG	1500
QY	1501	aaaaattctgcgcccaaggagactcgtatgaatcactctgtgagcaactttgcgcgaacaaagggataccc	1560
Db	1501	AAAAATCTTCCGCCCAAGGAGACTGATGGCATATCTGGCCAATTTTGCCCCGACACGGGGATCCC	1560
QY	1561	aattgagccccgaagaaccccaagaagggcccaacaatgtgccccctgtacaacgvcgvggtgtccagag	1620
Db	1561	AATTGAGCCCCGAAGACCCCAAAAGGGCCCAAAATGGCCCCGTTACAGGGGGGGGTCTACACAG	1620
QY	1621	taagttagttctgtgaactgtcgagcccgctgtgaagtgtcgvgcgvggtgtgtgtgtgtgtgtgtgt	1680
Db	1621	TAACTTAACTTGTGAACCTGTGGGCGCCCTGTGAGTGTGCGCGGGGGGTGTGCGGCCCAAGGCTTCC	1680
QY	1681	ggcctctcgaagacgctctctccccaataattgtctcagaacgacacttga 1725	
Db	1681	GCCCTCTGGAACCGCTTCTCCCAATAATGTCTACAGCGCCACTCTGA 1725	

RESULT	2
HIMACHIE	
LOCUS	
DEFINITION	HUMACHE 2218 bp mRNA
ACCESSION	Human acetylcholinesterase (ACHE) mRNA, complete cds.
VERSION	M55040
KEYWORDS	M55040.1 GI:177974
SOURCE	acetylcholinesterase.
ORGANISM	Human 21-week old fetus DNA, and cDNA to mRNA.
REFERENCE	Homo sapiens
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 2218)
JOURNAL	Soreq,H.E., Ben-Avitz,R., Prody,C.A., Seidman,S., Gnatt,A.,
MEDLINE	Neville,L., Biehan-Hurwitz,J., Dev-Dehman,E., Ginzberg,D.,
FEATURES	Leplidot-Blissom,Y. and Zakut,H.
	Molecular cloning and construction of the coding region for human
	acetylcholinesterase reveals a G + C-rich attenuating structure
	Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9688-9692 (1990)
	91088577
	Location/Qualifiers

FEATURES	LOCATION/QUALIFIERS
----------	---------------------

Location/qualifiers

```

source      1. .2218
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /map="7q22"
             /dev_stage="21-week-old fetus"
             /tissue_1lb="1ambda-gt10"
             1. .2218
             /gene="ACHE"
             /note="G00-118-746"
             /product="acetylcholinesterase"
             1. .2218
             /gene="ACHE"
             /EC_number="3.1.1.7"
             /EC="ACHE"
             /product="acetylcholinesterase"
             /codon_start=1
             /protein_id="AA68151.1"
             /db_xref="GI:177975"
             /db_xref="GDB:G00-118-746"
             /translation="MRPQCLHTPISLASPLLLMLLGGVGAEGHEDALLTVR
             GRLRGILKTPGVSALFGLPFAEPMGRRRLPPKQPMSCVNAVATFOSVCO
             YVDLYPGEETEMNPNRLESDCLYNTWTPYPRPSPIVLWITGGGFSASS
             LDVYDRETNQAEPTVLYSMNRYGAEFLALPGREARGNVGLDLQALQMGENT
             AAFGDPSTVTLFESGASVGMHLSPSSGLFRAVLDSGANGFWATVGMGEAR
             RRAQDLALVCCPGGTGNDTELVALRTPRADVLVHMHVLEPQESVFRSEPVV
             DGDPLSDTPEALINAGDFHGLQVLGVVKGDSYFLVYGAGFSKDNLSLSRAFLA
             GVRVGPVSDLAEAIVLHYTMDLHPEDPARLSDVYGDHNVCPVALQRLA
             AGARVAVYVEFHRASTLWPLMNGVPHGEIEEFGIPLDSRYVTAEEKFPAQRLM
             RYMANFATGDBNPRDRAKAPOMPRYTAGAOOYSLDPLPLEVRGLTAQACATPWR
             LPKLSATDITDEAEKRWKAEPHRSSTVMHKNDFHTSKDRCSDL"
             229. .291
             /gene="ACHE"
             /note="G00-118-746; does not fit consensus"
             310. .1998
             /gene="ACHE"
             /EC_number="3.1.1.7"
             /note="G00-118-746"
             /product="acetylcholinesterase"
BASE COUNT 355 a 755 c 680 g 428 t
ORIGIN
Query Match 99.8%; Score 1722; DB 9; Length 2218;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgaagcccccgagtgctgtctgcaacagccttcctgtctcccaatccttcctcc 60
DB 157 ATGAGGCCCCCGAGTGTCTGTGCAACAGCCTTCCCTGCTCCCACTCTTCCTC 216
QY 61 ctctctctgctcctgtgtgagagagtgtgaggtcgtgaagcgcgagagatgcaagctgtctg 120
DB 217 CTCCTCTGGCTCTCTGGTGTGAGAGAGTGGGGGCTGAAGGCGCGGAGATGACAGCTGCTG 276
QY 121 gtgaagtgctgtgagggcgagctgtgagcaltcgctgaagaaccccgaggggagcctgtc 180
DB 277 GTGAGCGTGTGCTGGGGGCGGCTGCGGCAATCCGCTGAAGACCCCGGGGCGCTGTGC 336
QY 181 tctgtcttcctggcattccctcttgagagcaccacatggagcccgctgtcttcctgcca 240
DB 337 TGTGCTTCTCTGGGCAATCCCTTGTGGAGCCACCATGAGACCCCGTGTCTGTGCA 396
QY 241 ccgagcccccaagcctctgtcagaggtgtgtagacgtacaaaccttcagagtgctctgc 300
DB 397 CGGAGGCCCAACAGCCTTGTCAGGGGTGTAGAGCTACAAACCTTCCAGAGTGTCTGCG 456
QY 301 taccaatatgtggaacacctataccaggtctttagagcaccagagatgtgaaaccccaac 360
DB 457 TACCAATATGTGGACACCTTATACCCAGGTTTGAAGGCAACCGAGATGTGAAACCCCAAC 516
QY 361 cgttgagctgaaggaagacgtcctgtacctcaacgtgtggaacccatacccccggctata 420

```

```

DB 517 CGTAGACTGAGGAGAGACTGCTGTACCTCAACGTTGTGCAACCAATACCCCGGCTACA 576
QY 421 tcccccacccctgtccctctgtgatactatagggggtgtctctacagtgtggcctctcc 480
DB 577 TCCCCACACCCGTCTCTGTGATCTATGGGGGTGGCTTACAGTGTGGGCTCTCTCC 636
QY 481 ttggaagctgaagatgagcgtctctgtgtacagagccgagagagactgtctgtgtctatg 540
DB 637 TTGAGAGCTGACAGATGGCCGCTTGTGTGTGACAGCGGAGAGAGCACTGTCTGTGTCA 696
QY 541 aactacccggtgtggagcctttgtctccctgtgcggccgtgcggagagccggagccggagc 600
DB 697 AACTACGGGTGGAGCCCTTGTGCTTCTGTGCGCTGTCCGGAGCCGAAGCCCGGGGC 756
QY 601 aatgtgtgtcctctgtatcagaaggtctgcccctgcaatgtggtcgaaggaagctgtgagcc 660
DB 757 AATGTGTGTCTCTGTATGAGAGCTGTGCGCTGTCACTGTGTGAGAGAACTGTGCAAGCC 816
QY 661 ttccggggtgacccgagacatcaatgtagcgtgttggggagagcgcgagagcgcctcgtgtg 720
DB 817 TTCCGGGGGTGACCCGCAATCAGTGAAGCTGTGTTGGGGAGAGCGCGGGAGCCCGCTGCTG 876
QY 721 ggcattcaactgtctgtcccgcccgccagccggggtcgtgttccaaagggcgtgtgtcagagac 780
DB 877 GGCATTCACTGTCTGTCCCGCCCAAGCCGGGCGCTGTCCACAGGGCGCTGTGCAAGAC 936
QY 781 ggtgcccccaatggaacctgtggcgcaagctgtggagatgagagagcccgctgcagagggcag 840
DB 937 GGTGCCCCCAATGAGACCTGTGGCGCAAGCTGTGCATGTGAGAGAGCCCGCTGCAAGGCCAG 996
QY 841 cagctgtgcccacctgtgtggcgtgtctctcagaagcgagcgtgtgtggaaatgaaacagaagcgtg 900
DB 997 CAGCTGTGCCCACCTTGTGGGCTGTCTCTCAAGCGCGCACTGTGTGGAATGACACAGAGCTG 1056
QY 901 gtagcctgtccttgagacagagacagagcaggtgtcgtgtgaaacagaatgagacgtgtg 960
DB 1057 GTAGCTGTGCTTGTGACAGAGCAAGCGCAAGTCTGTGTAAACCAAGAGGCAAGTGTG 1116
QY 961 cctcaagaagaagcgtctctcgtgtctctcgtgtcgtgtgtatgtagaactcctcagct 1020
DB 1117 CCTCAGAAAGGCTGTTCGGGTTCTCTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1176
QY 1021 gacacccagagagccctcattcaagcgagagacttccacagcctcgaagtgctgtgtgtgt 1080
DB 1177 GACACCCCAAGAGCCCTCATCAACGCGGAGAGACTTCCAGCGCTGTGAGGTGTGTGTGTGT 1236
QY 1081 ggtgtgaagatgagagcgtctgtattctcgtgtttaaaggggccccagagcttcagcaaaagc 1140
DB 1237 GTGTGAAGGATGAGGGCTGTGTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1296
QY 1141 aacgaagctctcattcaagcgagagcttccctgtgcccgggtgtgaggtctgaggtctcccaag 1200
DB 1297 AACGAGTCTCTATCAAGCGCGGCGAGTCTCTGTGCGGGGTGTGCGGGGTGTGCGGGGTGTGCG 1356
QY 1201 gtaagtgaacctgtgacgagcagcgtgtgtctcgtgalttaacagaactgtgtgtgtgtgtgt 1260
DB 1357 GTAAGTGACCTGTGCAAGCCAGCGCTGTGTCTCATTTACACAGATGCGTGTGTGTGTGTGTGT 1416
QY 1261 gaccggagcagcctgtgagagagcctcgtgagcagatgtgtgtgtgtgtgtgtgtgtgtgtgt 1320
DB 1417 GACCGGCAAGCGCTGTGAGGGAGGCGCTGTGAGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGT 1476
QY 1321 cccgtgtgccaagctgtgctgtgagcagcgtgtgcccaggtgtgcccaggtctcagctcgtc 1380
DB 1477 CCGGTGTGCGCAAGCTGTGCGGCGACATGTGCGTGTGCGGAGGTGTGCGGCGGTGTACGCTTACG 1536
QY 1381 ttggaacccgtgtgtcagcagctctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1440
DB 1537 TTTGAACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1596
QY 1441 ggaatcagatcaactcttttgatccccctgtgagccctctcgaatacacaacgagagagagag 1500
DB 1597 GAGATGAGATTCAATCTTTGGGATCCCGCTGTGAGCCCTGTGTGAATCAACGGGAGAGAGAG 1656

```

QY 1501 aaatcttcgcccagcagctgatacgatctgagccaaatttgcgcgacagggagatccc 1560
|||||
Db 1657 AAAATCTTCGCCGACGACACTGATGATACAGGGCCAACTTTCGCCGACGACGAGATCCC 1716
|||||
QY 1561 aatgaagcccccagagcccccagagcccaaatgagcccccagatgagcccccagggggagcagag 1620
|||||
Db 1717 AATGAGCCCGACAGACCCCAAGCCCAATGAGCCCGGATACAGCGGGGGCTCAGCAG 1776
|||||
QY 1621 taagttagctgagacgtcgagccgctgagaggtcgagcgagggagctgcgcgcccagagcctgc 1680
|||||
Db 1777 TACGTTAGTGTGAGACGTGCGGCGCCTGTGAGAGTGGGGGGGGCTGCGGGCCCAAGCCTGC 1836
|||||
QY 1681 ggccttcggaacgcgctcctcctcccaattgtcagagccacc 1722
|||||
Db 1837 GCGTTCGTGGAACCGCTTCCTCCCAATTCCTCAGCGCCAC 1878
|||||

RESULT 3
AR070205 AR070205 2256 bp DNA Linear PAT 18-FEB-2000
LOCUS Sequence 5 from patent US 5891725.
DEFINITION AR070205
ACCESSION AR070205
VERSION AR070205.1 GI:7221093
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2256)
AUTHORS Soreq, H., Zakut, H. and Eckstein, F.
TITLE Synthetic antisense oligodeoxynucleotides and pharmaceutical
JOURNAL compositions containing them
FEATURES Patent: US 5891725-A 5 06-APR-1999;
source Location/Qualifiers
1..2256
BASE COUNT 390 a 757 c 680 g 429 t
ORIGIN

Query Match 99.8%; Score 1722; DB 6; Length 2256;
Best Local Similarity 100.0%; Pred.-No: 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagagcccccagctgtctgtctgacacagccttcctgtcctcccaatcctctctc 60
|||||
Db 160 ATGAGGCCCGCCGACGTGCTGCTGACACAGCGCTTCCTGCGCTCCACATCTCTCTC 219
|||||
QY 61 ctctctgtgctcctgagtgagagagtgaggcttgaagcccgagagagtgagagctgtg 120
|||||
Db 220 CTCTCTGTGCTCTGTGGTGAGAGTGGGGCTGAGGGCCGGAGAGTGCAGCTGCTG 279
|||||
QY 121 gtgagagtgagtgagagcgagctgagagcattctgcctgaagaccccgagggcgctgtc 180
|||||
Db 280 GTAGAGGTGGTGGGGCGGCTGCGGCAATTCGCTGAAGACCCCGGGGGCGCTGTC 339
|||||
QY 181 tctgtcttcctgtgagccttcctgtgagagccacccatgtgagcccgctgccttcgtca 240
|||||
Db 340 TCTGCTTCCGAGGCAATCCCTTTGCGAGCACACCATGGAGCCCGCTGCTTCTGCA 399
|||||
QY 241 ccgagagcccaagcagcctgtgtcaagggtgtgtgagcgtacaaacttccagagtgctgc 300
|||||
Db 400 CCGAGGCCCAAGCCCTTGTGGTCAAGGGGTGTAGACCTTCAACCTTCCAGAGTGTGTC 459
|||||
QY 301 taacaatactgtgacacccataaccaggttttgaaggacccagagatgtgaaccccaac 360
|||||
Db 460 TACCAATATGTGACACCTTATCCCAAGTTTTGAAGGACCGAGATGTGAACCCCAAC 519
|||||
QY 361 cgtgagctgagcgagagcgtcgtgtacccaacgtgtgtgacacccaatacccccgtctaca 420
|||||
Db 520 CGTAGCTGAGCGAGGAGACTGCTGTACCTCAACGTGTGACACATCCCGCGGCTACAA 579
|||||
QY 421 tcccccacccctgtccctcgtctgtgatactatgtgggtgtgcttctcaagtgtggcctctcc 480
|||||

Db 580 TCCCCACCCCTGTCTCTGTGATCTATGGGGGTGCTTCTACAGTGGGGCTCTCC 639
|||||
QY 481 ttgagcgtatacgtatgagccttcttggtaagagccagagagctgtgtgtctcatg 540
|||||
Db 640 TTGGACGTGTACATGAGCGGCTTCTGTGTACAGCCAGAGAGATGTGTGTCTCATG 699
|||||
QY 541 aactacccggttggagccttgtctcctgtcgcctgcgagagccgaagagcccgagc 600
|||||
Db 700 AACTACCGGCTGGAGAGCTTGTGCTCTGCTGCTGCGCGGAGACCCAGAGCCCGCGC 759
|||||
QY 601 aatgtggtctcctgtatcagaagcttgccctgagtggtgtgagaggaagtgtgagcc 660
|||||
Db 760 AATGTGGCTCTCTGTGATCAGAGCTGCTGCTGAGTGTGTGCAAGAAAGTGTGAGCC 819
|||||
QY 661 ttccgggttgaacccagacatcagtgacgtcttgggtgagagcgagcgagcctcgtgtg 720
|||||
Db 820 TTCGGGGGTGACCCGACATCATGACGCTTGTGGGAGAGAGCGGGAGCCGCTCGTG 879
|||||
QY 721 ggcatacactgtgtctcccgccagccgggagcctgttccaaagggccgtgtcagaagc 780
|||||
Db 880 GGCATGCACTGTCTGCTCCCGCCAGCGGGGCTGTTCCAGAGGCGCTGCTCAGAGC 939
|||||
QY 781 ggtgcccccagatggagccttgagcagagtgagcattggaagagccgttcgagggccagc 840
|||||
Db 940 GGTGCCCCCAATGAGACCTGGGCCAGGGTGGCAATGGAGAGGCCCGTTCGACGGGCCAGC 999
|||||
QY 841 cagcttgagccacactgtgtgtgtctcctcagcgagcagctgtgtggaatgaacagagctg 900
|||||
Db 1000 CAGCTGGGCCACCTGTGTGGCTGTCTCTACAGCGGACATGTGGGAATGACACAGAGCTG 1059
|||||
QY 901 gtgagcgtccttcgagacagacagcagcgagcagctgtgtgaacacgaatgtgacagctgtg 960
|||||
Db 1060 GTACCTGCTCTTCGAGACACGACAGCGCAGCTCTGTGTGAACACGATGCGACGTCTG 1119
|||||
QY 961 cctaaagaaagcgtcttcctcctgtctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
|||||
Db 1120 CTTCAAGAAAGCTCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
|||||
QY 1021 gacaccccaagagccctcatcaacagcgagagcttccacagcctgcagctgtcgtgtgt 1080
|||||
Db 1180 GACACCCCAAGAGGGCTCTCATTAACGCGGAGACTTCCAGCGGCTGTGAGTCTGTGTG 1239
|||||
QY 1081 gtgtgtgaagagtgagagcgtgtatcttcgtgttcaagggccccagagcttcaagaaagc 1140
|||||
Db 1240 GTGTGTGAAGGATGAGGGCTGTATTTCTGTGTACGGGGCCCAAGGCTTACGCAAAAGAC 1299
|||||
QY 1141 aagagctctcatcagccgggagagcttctgtgcgggggtgtgaggtgtgcgggttcccgag 1200
|||||
Db 1300 AAGAGTCTCTCATCAAGCGGGGCGAGTCTGTGGCCGGGGTGTGGGGTGTCCCGAG 1359
|||||
QY 1201 gtaagtgaactgtgacagcagagctgtgtcttcattacaacagactgtgtcattcccgag 1260
|||||
Db 1360 GTAAGTGAACCTGTGAGCGAGAGCGTGTGTCTGTGATTAACAGACTGTGTCAATCCGAG 1419
|||||
QY 1261 gaaccggcagcgtgagagagccctgaagcgaatgtgtgtgagagccacaatgtcgtgtgc 1320
|||||
Db 1420 GACCCGGCAGCCTGTGAGGAGAGCGCTGAGCATGTGTGTGGCGACCAATGTGTGTGTC 1479
|||||
QY 1321 ccggtggcccaagctgtgtgtgagcgtgtgtgtcccaaggtgtccgggtgtcctacgtatgtc 1380
|||||
Db 1480 CCGGTGGCCCAAGCTGTGGGTGGCGAGTGTGCTCCAGAGGTGTCCCGGTGTAGCGCTTACGTC 1539
|||||
QY 1381 tttaaacacgctgtctcacaagcctcctcgtgagccttgtgtgtgtgtgtgtgtgtgtgtgt 1440
|||||
Db 1540 TTTAAACACGCTGTCTCACGCTCTCTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1599
|||||
QY 1441 gagaatgagatcatcttgttgatcccccgtgaaccccttcgaacatacaagcgagagag 1500
|||||
Db 1600 GAGATGAGTTCATCTTGTGGATCCCTGTGAGCCCTCTGAAACTACAGCGCAGAGAGAG 1659
|||||
QY 1501 aaatcttcgcccagcagcagtgatgatactgtgagcaatttgcgcgacagggagatccc 1560
|||||

Accession	Sequence	Position
Db	AAAATCTTGCCCAACGACTGATGGAGATACAGGGGCAACTTTGGCCGACAGGGGATCCC	1719
Qy	1551 aatgagcccccagagaccaccaagggccccaacatgagcccccgtacacgycgggagctcagcag	1620
Db	1720 AATGAGCCCCCGAGACCCCAAGGCCCAACATGAGCCCCCGTACACGGCGGGGCTCAGCAG	1779
Qy	1621 tacgttaagtctcggaaacctcggcgccgctgtaagtgtcgycgggagctctcgcgcccagagcttgc	1680
Db	1780 TACGTTAAGTCTCGAGACTCGGGCGCTGGAGAGTGGCGGGGGGCTCGCGGCCACGAGCCTTC	1839
Qy	1681 gcccttcgaaacgcgtctctcccacaattgctcagaagcacc	1722
Db	1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATGTCTACACGGCGACC	1881

RESULT	4				
AR070207					
LOCUS	AR070207		3016 bp	DNA	linear
DEFINITION	Sequence	7 from patent US 5891725.			
ACCESSION	AR070207				
VERSION	AR070207.1	GI:7221095			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				

REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 3016)
TITLE	Soreq, H., zakut, H. and Eckstein, F.
JOURNAL	Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them
FEATURES	Patent: US 5891725-A 7 06-Apr-1999;
source	Location/Qualifiers 1. .3016

	/organism="unknown"			
BASE COUNT	497	a	1065	c
ORIGIN	840	g	614	t

```
Query Match          99.8%; Score 1722; DB 6; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY	1	a t g a g g c c c c c a g t g t c t g t a c a a c g c t t c c c t g c t t c c c a c t c t t c t c t c	60
Db	160	A T G A G G C C C C C C G A C G T G T G C T G C A C A C G C C T T C C C H G C C A C T C T T T C C T C	219
OY	61	c t c c c t t g c t c t c t g t g t g y a g a b l t g g y g c t t g z g g c c g y g a a g a r g c a a g t c t g t	120
Db	220	C T C C T C T G G C T C T C T G G G T G G A G A G T G G A G G C T G A G G C C G G G A G A G T C A A G T G C T G	279
OY	121	g t g a c g t g c g t l g y g g c c g y c t g c y g g g a t l c g c c f y a a g a c c c c g y g g c c t g t c	180
Db	280	G T G A C G G T G C G H G G G G G C G G C T G C G G G G C A T T G C C T G A A G A N C C C G G G G G C C T G T C	339
OY	181	t c t g t t t c c c g y g g a t c c c c t t g c y g a g c a c c c a t g y g a c c c c g t c o g t t t c t g c a	240
Db	340	T C T G C T T T C T C T G G G G A T C C C T T T T G C G A G C A C C A T C A T G G G A C C C C G T G C T T T T G C C A	399
OY	241	c o g y a g c c a a g c a a c c t t g t c a a g y g t g y t a a c g c t t a c a a c c t t c a a g t l c t g c	300
Db	400	C O G G A G C C C A A G C A C C T T M G G T C A G G G T G T A G A C G C T A C A C C T T C C A G A G T G T C T G C	459
OY	301	t a c c a a t t t g t g y a a c a c c c t a t a c c a a g t t t t g a y g y a c c g a g a t g t y g a a c c c c a a c	360
Db	460	T A C C A A T T T G T G A A C A C C C T A T A C C C A G A G T T T T G A G G G C A C C G A A T G T G A N A C C C A A C	519
OY	361	c g t g a c t g a g a g a g a c t g c t g a c t t a c c t c a a c g t g t g a c a c a t a c c c c g y g c t t a a	420
Db	520	C G T G A G C T G A G A G A G A C T G C C T T A C C T A A C A C G T G T G A C A C A T A C C T C C C G G G C T F A C A	579
OY	421	t c c c c a c c c c t g t c t c t g t c t g a l c t a t a y g y g t g y c t t c a g t y g g a g c t c t c t c c	480
Db	580	T C C C C A C C C C C T G T C T C T G T G A T C T A T A G G G G T G G C C T T C T P A A G T G G G C C T C T C C C	639

QY	441	ttgtagcgtgtacgaatgacgacgtctctcttggtacagagccgagaaagacgtgtcgtgtacatg	540
Db	640	TTGGACAGGTGACGATGGCCGCTTCTTGTGTACAGAGCCGAGAGAGACTGTGCTGTCTCATG	699
QY	541	aaccaccggtgtgagagaccccttggctctccctggccctgcgaggagaccgagagccccgggc	600
Db	700	AACATACCGGGGTGGAGACCTTTTGGCTTCTGTGCTGTCCGGAGCCGAGAGGCCCCCGGC	759
QY	601	aaatgtgagctctctctggaatcagaaagctgagccctgagctgagtgagtgacagtgacacg	660
Db	760	AATGTGGGTCTCTGATATCAGAGGCTGTGGCCCTCATGTGCTGACAGAGACGTGTGCACGC	819
QY	661	ttcgagggtgtgaccccgaaacaatgtacgctgttttgaggagagcgaggagccgctgggtg	720
Db	820	TTCCGGGGGTGACCCGACATATGTAAGACTGTTTGGGGAAGAGCGGAGCCGCTCTCGATG	879
QY	721	ggcatgcaacctgcgtctcccccagaccgaggagccgtctcacaagggccgtctgtagaagc	780
Db	880	GGCATGCAACCTGCTGTCTCCCGCCACGCGGGCTGTTCACAGGGCCGTGTCTGCAGAGC	939
QY	781	ggtgcccccaatgagacccctgggacaggtgagcatgagaaagaccctgcagaggccacg	840
Db	940	GGTGCCCCCAATGAGACCTTGGCGCACAGGTTGGCATGGAGAGAGCCGTTGCGAGGGCCAGC	999
QY	841	cagctgtgccccacccctgtgtgggctgtctctccagcgcgcaactcgtgtgggaatgacaagactg	900
Db	1000	CAGCTGTGCCCCACCTTGTGTGGGCTGTCTCTCCAGGGGCACTGTGTGGGAATGACACAGACTG	1059
QY	901	gtagccttgaccttgagacaagacagcgcaagctccgtgtgaacacagaatggaacgtgtg	960
Db	1060	GTAGCCTTGACCTTGCAGACACACAGCGCATGCTGTGTGAACCGAATGGACGACTGTGTG	1119
QY	961	ccctcaagaaagcgtctctccggtctctcccttgctgcctgtgtagttagtgaagactctccacgt	1020
Db	1120	CCTCAAGAAAGCGTCTTCCGGGTTCTCTCTCTGTGCTGTGTGTGATGAGACTTCTCTAGT	1179
QY	1021	gacaccccaagagccctcatcacaacgaggagactctcacgacctgcaggtgtcgtgtgtgt	1080
Db	1180	GACACCCCAAGAGGCCCTCATCAACAGGGGAGACTTCACAGGCTCGAGTGTCTGTGTGGGT	1239
QY	1081	gtgtgtgaagagatgtagggctcgtatcttctcgtgtttagagggcccgagagcttcagaaagac	1140
Db	1240	GTTGTGAAGATGAGGGCTCTGTATTTTCTGTGTGTTTACGGGGCCCAAGGCTCATGCAAAATAC	1299
QY	1141	aacgagctctcatcaacagccgggacggagcttccgtgcgggggtgcgggtctcgaggtctccacg	1200
Db	1300	AACGAGTCTCTCATACAGCCGGGCGAGATTCTGTGGCGGGGTGCGGGTGTGCGGTTCCGAG	1359
QY	1201	gtaaagtgtacgtgcagccgagagcgtgtgtgtctgtcatctacacagactgtgtcatcccgag	1260
Db	1360	GTAAGTGAACCTGTGGACGCGAGGGCTGTGTGTCTCTCATTAACAGACTGTGCTGATCCGAG	1419
QY	1261	gacccgagaaacgctgtgaggagggcccttgagagatgtgtgtggagacaacaatgtctgtgc	1320
Db	1420	GACCCGGACACCTTGAGAGGAGGGCCTGTAGGAGTGTGTGGGACACACATGTCTGTGTC	1479
QY	1321	cccgctgagcccaagctgtgcctgaggcagactgtctgcacaaggtgtcccgaggtctacagctac	1380
Db	1480	CCCGTGGGCGAAGCTGTGGGCGACTGTGGCTGCCACAGAGTGCCTCGACTACACCTTACACTC	1539
QY	1381	tttgaacaacgctgtctccacgctctccctgtgcctgttgaatgagggtgtgcgcccaagctac	1440
Db	1540	TTTGAACAACCTGTCTTCCAGAGCTCTCTGTGCCCCCTTGTGATGGGGGTGCCCAACGAGCTAC	1599
QY	1441	gagatgagttcatcatttggatctcccccctggagacccctcgaaacttaaacgacgaagag	1500
Db	1600	GAGATGAGTTCACTTTTGGGATTCCTCCCTGTGAGACCTCTGTGAAATACACGGCAAGAGAG	1659
QY	1501	aaaatcttcgcccagcgacatgacgtacatctgagccaactttggccgcacagaggagatccc	1560
Db	1660	AAAATCTTTCGCGCGAGCACTGATGTGCAATCTGTGGCCAACTTTGCCCGCACAGGGATATCC	1719
QY	1561	aaatgagccccagagacccccaagacccccaatgtgccccctacacggtcgagggtctcagag	1620

Db	1720	AATGAGCCCGGAGACCCCAAGGCCCCACAAATGAGGCCCGCTACAGCGGGGGCTCAGCAG	1779
Qy	1621	taacgttagctctggaacctgcgcgcctgtgaagtgcggcggggtctgcgcgcgcgcgcctg	1680
Db	1780	TACGTTAATCTGGAGACCTCGCGCCTGGAGGTCGCGCGGGGCTCGCGCCAGCCCTGC	1839
Qy	1681	gaccttggaaacgcttccctcccaattgctcaagcgccacc	1722
Db	1840	GCCCTTCTGGAAACCGCTTCTCTCCCAAAATTGCTCACGCCGCACC	1881
RESULT	5		
LOCUS	AR070206	3096 bp	DNA
DEFINITION	Sequence 6 from patent US 5891725.		linear
ACCESSION	AR070206		
VERSION	AR070206.1		
KEYWORDS	GI:7221094		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 3096)		
AUTHORS	Soreq,H., Zakut,H. and Eckstein,F.		
TITLE	Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them		
JOURNAL	Patent: US 5891725-A 6 06-APR-1999;		
FEATURES	Location/Qualifiers		
source	1..3096		
BASE COUNT	509 a 1089 c 872 g 626 t		
ORIGIN	/organism="unknown"		
Query Match	99.8%; Score 1722; DB 6; Length 3096;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 1722:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	atgagggcccgagatgctctgcagacacgccttcctgcttcccactcttctctc	60
Db	160	ANGAGGCCCGCCAGCATGTCGTGCACACAGCCCTTCCCTGCTCCCTC	219
Qy	61	ctccctctgctctctgtgtggaagatgggggctggaagccggagagatgacaagctgtc	120
Db	220	CTCCTCTGGCTCTTGTGGTGAAGACTGGGGCTGAGGGCCGGAGAGTCAAGCTGCTG	279
Qy	121	gtgaagcgtgtggtgggacgcggtctgcgggacattgccttgaagaccgccggggacctgtc	180
Db	280	GTGAGCGTGTGGGGGCGCGGCTGCGGGGCATTCGCTGAAAGACCCCGGGGGCCCTGTC	339
Qy	181	tctgcttctctgcgcatccctcttgcggaagcccaatggaaacccgcttcgttcgca	240
Db	340	TCTGCTTCTCTGGGCAATCCCTTTTCGGAGCACCAATGAGACCCCGCTTCTGCCA	399
Qy	241	ccggagcccaagacaccttgtcaagggtgtgtagacgttacaaccttcagaagtctgc	300
Db	400	CCGGAGCCCAACACACCTTGTGCAAGGGGTGAGACGCTACAACTTCCAGAGTCTGC	459
Qy	301	taccataatgtgagacacctatacccaaggttttgaaggcaaccgaatgtgaaccacaac	360
Db	460	TACCAATATGTGAGACACCTATACCAGGTTTGAAGGCGACCGAATGTGGAACCCCAAC	519
Qy	361	cgtgagcgtgagcgagactgcctgtaacctcaacgltgtgaacacatacccccgactaca	420
Db	520	CGTGAAGTGAAGAGAGACGACCTGTGACCAACGHTGAGACACATACCCCGGCTGACA	579
Qy	421	tcccccacctgtctcgtctgaltcaltabvggggtgagcttctacagtggggctctcc	480
Db	580	TCCCCACACCCCTGCTCTGTGTGATCATATGGGGGTGGCTTCTACAGTGGGGCTCTCC	639
Qy	481	ttggacggttaagatggcgccgtctctctgtgtacaagcgaggaagacgttgtgtgcatg	540
Db	640	TTGGACGGTTAAGATGGCGGCTTCTTGTGATACAGCGCGAGAGGACTGTGTGTTCATG	699

Oy	541	aaacacgggttgaggaagccttttgctcttcacggcccttgccggggagacgcggagagcccccggc	600
Db	700	AACtACCGGGGTGGGAGCCTTTTGGCTTCtTGcCCtTGccGGGAGCCGAGAGCCCCGGCC	755
Oy	601	aatttggtctcctbgtatctcaagagcttgcccttgacgttggtgtcaagagaaacgttgacgc	660
Db	760	AATGTGGGTCTCCtGTATCAGAGGCTTGCCCTCTGcATGTGGTGTCAAGAAAGTGGAGCC	819
Oy	661	ttcgggggtgacccggaatactagtagcagctgttttgggagagcgcgggagccgcttggtg	720
Db	820	TTCCGGGGGTGACCCGAGATCATGTGAGCGCTTTTGGGGAGAGCGCGGAGCCGCTCGGTG	879
Oy	721	ggaatgcaacctgctgttcccgccacgcggggccgtgttccaagggccgttgctcaagc	780
Db	880	GGCATGCACTCGCTGTCTCCCGCCACCGGGCCCTTTTCCAGAGGCCGCTGTCTGACAGC	939
Oy	781	ggtgcccccgaatbgaaaccttgggccaagcttgaggcaatggagagagcccgctgcagggccacg	840
Db	940	GGTGCCTCCCAATGGAACCTTGCGGCACAGGTGGGCGATGGAGAGGCCCCGTGCAGGGGACAG	999
Oy	841	cagcttggtgccacacttggtggctgttcctccaagccgagcactggtgggaaatgacaagaagctg	900
Db	1000	CAGTGTGGCCCACTTGtGGGCTGTCTCTCCAGGGGGGACTGTGGGAAATGACACAGAGCTG	1059
Oy	901	gtagcccttgcttcgagacaacacacgacgacgagtgcttggttgaaacaagaaatggcaagctgtg	960
Db	1060	GTACCTGTGCTTGCGAGACAGCAAGGAGAGGTCTGTGTAAACAGAAATGGCCAGCTGTG	1119
Oy	961	ccctcaagaagaagcgtcttcoggtcttccttcctgcgttggtgtgtatagatgaaagactcccaagt	1020
Db	1120	CCTCAAAAGAAAGCGTCTTCCTCCGGTCTCTCTTCGTGCTGTGTGTGAATGGAGACTTCCTCAGT	1179
Oy	1021	gaacccccaagagccctccatcaaacgcggggagagacttccaaagcctgtcaagtgctgtgtgtgt	1080
Db	1180	GACACCCCAAGAGGCCCTCATTAACGGCGGAGACTTCCAGGCCCTCAGAGTCTGTGTGGGT	1239
Oy	1081	gttggtgaagagatagaggctctgtaattcttggtttaaagggccccacagccttcagcaaaagac	1140
Db	1240	GTGGGAAGGATGAGGGCTGTGTATTTTTCtTGTTTACGGGGCCCCAGCGCTTCAGCAAAAGAC	1299
Oy	1141	aaagagctctcatcaagccggggccagagcttccttgccgggggtgcgggttcgggggtctcccgag	1200
Db	1300	AACAGAGTCTCTCATACACCCCGGCCGAGTTCTGTGGCCGGGGGTTCGGGGTTCGCCAG	1359
Oy	1201	gttaggtaccctggagagcccgagcgtctgtgtccctgcctatcaaaagaaatggctgcatacccgag	1260
Db	1360	GTAAGTGACTTGCGACGCCGAGGCGTGTGTGTCtTGcATTACAGAGACTGGCTGcATGCCGAG	1419
Oy	1261	gaaccggcacagcctbtaaggagagccctgtagcgaatgtgttggtgcacacaatgtcgtgtgc	1320
Db	1420	GACCCGGACGCGCTGAGGGAGGGGCCCTTGAGCGAGTGTGTGGGGACACAAATGTGCTGTCG	1479
Oy	1321	cccggtggcccaagctggtcttggtggagactgctctgcccaagggtgtgcccggtctaaagctaaagtc	1380
Db	1480	CCCGTGGGCCAGCGTGTGGGCTGAGCTGTGCTGCCACAGGAGTCCCGGGGTGTACGCTTACGTC	1539
Oy	1381	tttgaacaacgctgtcttccaagctctccctcgccccctgtgatatgggtgtgtccccaagcgtac	1440
Db	1540	TTTGAACAACGCTGCTTCCAGCGTCTCTCTGCCCCCTGTGTGATGGGGGTGCCCCACAGGCTAC	1599
Oy	1441	gagaatcagatcatalcttttggatctccccccttggaaacccctctcgaaactaaacgtacagagagag	1500
Db	1600	GAGATCGAGTTCAATCTTTTGGGATTCGCCCTGAGACCCCTTCGAAATACACGGCAGAGAGAG	1659
Oy	1501	aaaatctctgcacagagcactgtagtagaatacttggtgccaacttgcgcgcacaaagggatctcc	1560
Db	1660	AAAATCTTTCGCCACGACACAGATGCATACTGTGGCCAACTTGTGCCCGACAGGGGATCTCC	1719
Oy	1561	aattgaagcccccgaagcccccaagagcccaacaattggcccccggtaaagcggggggtctcaagag	1620
Db	1720	AATGAGCCCCGAGACCCCAAGGCCCAACATATGGCCCCCGCTACACGGCGGGGCTTCAGCAG	1779

Db 2404 CCCCCAGACCCCAAGGCCCAACAATGGCCCCCTACACGGGGGGCTCAGACGTACGTT 2463

QY 1627 agctcgaacctcgcgcgcctcgaagatgctcgccggcgccgaagcctcgccttc 1686

Db 2464 AGCTGAGACCTGGCGCGCGTGGAGAGTGGCGGGGGCTGGCGCCAGAGCGCTCGCCCTTC 2523

QY 1687 tggagaccgttcctccccaattgctcagcgc 1718

Db 2524 TGGACCGCTTCCTCCCAATGGTCTCAGCGC 2555

RESULT 7
AX275255 14446 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 4 from Patent WO0171014.
ACCESSION AX275255
VERSION AX275255.1 GI:16547675

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (sites)
AUTHORS Mor.T., Soreq, H., Arintzen, C. and Mason, H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants

JOURNAL
Patent: WO 0171014-A 4 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US); Mor,
Tsafirir (US); Soreq, Hermona (IL); Arintzen, Charles (US); Mason,
Hugh S. (US)

FEATURES
source 1..14446
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM036"

BASE COUNT 3231 a 3831 c 4046 g 3042 t 296 others

ORIGIN

Query Match 99.2%; Score 1712; DB 6; Length 14446;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ccccgcaatgctcgtcgaacagccttcctcgtcctccactcttcctccctc 66

Db 148 CCCCCAGAGTGTGCTGCAACAGCCCTTCCTGCTCCCTCCCTCTCTCTCTC 207

QY 67 tggctcctggctggaagagtgagggcgctcgaagagcggagatgacagctcgtcgtgagc 126

Db 208 TGGCTCTGCTGGTGGAGAGTGGGGGCTGAAGCGCGGAGAGATGACAGCTGCTGTACG 267

QY 127 gtagcgtggggcgccgctcgcgggacatcgctgaagaaccccgggggccctgctcgt 186

Db 268 GTGCTGTGGGGCGCGGTGGGGGCTGATTCGCTGAAAGACCCCGGGGCGCTGTCTGCT 327

QY 187 ttccctggagatccctcttcgagagcaccatggaaccccgctcgtcttcgcaacggag 246

Db 328 TTCTCTGGGCAATCCCTTTGGGGAGCAACCATGGAGACCCCGCTTCTGCAACGGAG 387

QY 247 cccaagcagcctcgtcaggggtgtagagcgtacaaacttcagagtgctcgtacaa 306

Db 388 CCAAGCAGCCTTGTGTCAGGGGTGTAGACGCTACAACTTCAGAGTCTCTACCAA 447

QY 307 tatgtgaacacctatacccaagctttagaggaacagagatgtgaaccccaacgtgag 366

Db 448 TATGTGAGACACCTTATACCAAGGTTTGAAGGGACGAGATGTGGAACCCCAACCGTAG 507

QY 367 ctgaagagagagcgtcgtacccaagcgtgtgaacacatacccccggcgctacatcccc 426

Db 508 CTGAGGGAGAGACTGCTGTACCTCAACGCTGTGACACCAATACCCCGCGCTACATCCCC 567

QY 427 acccctcctcgtcgtcgaactatcgggggtgagcttcacagtgaggcctcctcgtgag 486

Db 568 ACCCGTCTCTCTCTCTGAGATCTATGGGGGTGGCTTACAGTGGGGCTCTCTCTTGAG 627

QY 487 gctgaagatggcgcctctcttgtaacagggcgagagagctgtcgtgtgtccatgaactac 546

Db 628 GTGTAGATGGCGCTCTTGTGTACAGGCGGAGAGAGCTGTGTGTGTCTCAATAC 687

QY 547 cgggtggagaccttcttgcttcctcgcgcgcctcgggggagggcgccggcgcaatgtg 606

Db 688 GGGGTGGAGCCCTTTGGCTCTCTTGCCCTTGCGGGAGCCGAGAGGCTCCGGGCAATGTG 747

QY 607 gctcctcgtatcagagagctcgcctcgaatggatggagagagagcttgagccttcggg 666

Db 748 GGTCTCTGTATCAGAGGCTGGCCCTGCTGAGTGGGTGAGAGAGAGAGAGCTTTCGGG 807

QY 667 ggtgaaccgaacatcagtgagcgtgtcttgaggagagcgcgggagcgcctcgtgtgagc 726

Db 808 GGTGACCCGACATCAGTACAGTGTGGGAGAGCCGCGAGACCCGCTCGTGGGAGTGC 867

QY 727 cactcgtcgtcccgcccgagcgaggcggtcttcacagagcgctgtcgaagagcggtgc 786

Db 868 CACCTGCTGCCCCCGCCACCGGGGCTGTTCACAGGGCGCTGTGACAGCGGTGCC 927

QY 787 cccaatggaaccttgaggcagagtgaggatggagagcgccgtcgaagggcagcgaagctg 846

Db 928 CCAATGAGACCTTGCGCAGAGGTGGGATGAGAGAGCCCGTGGAGGCGCAGCAGCTG 987

QY 847 gcccaactgttggtcgtcctcgaagcgagcgtgtgtggaatgaacagagctgtgaagc 906

Db 988 GCCCACCCTTGTGGCTGTCTCTCCAGCGGCACTGTGGGATGACAGAGCTGTAGCC 1047

QY 907 tgccttcgagagagagcagagcagagctcgtgtgaacagagagagcaggtgcgcctcaa 966

Db 1048 TGCTCTGAGACAGACAGCAGCGAGAGCTCTGTGAACCAATGAGCAGTGTGCTCTCAA 1107

QY 967 gaaagcgtctcgcgtctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1026

Db 1108 GAAAGCGCTTCCGGGTCTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1167

QY 1027 ccaagggcctcctcaacagcgagagagctcgaagcgtcgtcgtcgtcgtcgtcgtcgt 1086

Db 1168 CCAAGAGCCCTCATCAACAGCGGAGAGCTTCAAGGCTGTGAGGTGTGTGTGTGTGT 1227

QY 1087 aagagtagagagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1146

Db 1228 AAGGATGAGGCTGTGATTTTCTGTTTACGGGCCCCAGGCTTCAACAGACAGAGAG 1287

QY 1147 tctcctacacagcgagagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1206

Db 1288 TCTCTATCAAGCGCGGCGAGTCTGTGCGGGGTGCGGGTCCCAAGTAACT 1347

QY 1207 gaactgagcagcagagctgtgtcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1266

Db 1348 GACTGTGACACCCAGAGGT 1407

QY 1267 gcaagcctgagagagagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1326

Db 1408 GCAGCCTGAGGAGAGGCGCTGTGAGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1467

QY 1327 gcccaagctgagctgagcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1386

Db 1468 GCCAGCTGCTGGCGAGCTGTGCGCCAGAGGTGCCCGGTTCAGGCTACGCTTTGAA 1527

QY 1387 caacgtgctcgaagctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1446

Db 1528 CACCTGTCTTCAACGCTTCTCTGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1587

QY 1447 gaggtaactcttgagatcccccgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1506

Db 1588 GAGTTCATCTTTGGATCCCTCTGAGACCTCTCTGAACTACAGCGGAGAGAAATATC 1647

QY 1507 ttcgccaagcagatgagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1566

Db 1648 TTGCGCCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1707

QY	1567	CCCCGAGAGCCCAAGAGCCCAACAATGGAGCCCGGTACAGAGGAGCTCAGAGAGT	1626
Db	1708	CCCCGAGAGCCCAAGAGCCCAACAATGGAGCCCGGTACAGAGGAGCTCAGAGAGT	1767
QY	1627	AGTCGTGACACTGCGGCGCGTGGAGGTGGGCGGAGGCTGCGGCGCCAGGCGTTC	1827
Db	1768	AGTCGTGACACTGCGGCGCGTGGAGGTGGGCGGAGGCTGCGGCGCCAGGCGTTC	1827
QY	1687	TGGAAACCGCTTCCTCCCAATGGCTGACAGCG	1859
Db	1828	TGGAAACCGCTTCCTCCCAATGGCTGACAGCG	1859
RESULT	8		
LOCUS	HUMACHE01		
DEFINITION	Human acetylcholinesterase (ACHE) gene, exon 2.		
ACCESSION	L22559		
VERSION	L22559.1 GI:853589		
KEYWORDS	acetylcholinesterase.		
SEGMENT	1 of 4		
SOURCE	Homo sapiens blood DNA.		
ORGANISM	Homo sapiens		
REFERENCE	1 (sites)		
AUTHORS	Soreq, H.E., Ben-Aziz, R., Prody, C.A., Seidman, S., Gnatt, A., Niville, L., Lileman-Hurwitz, J., Lev-Lehman, E., Ginzberg, D., Lapidot-Lifson, Y. and Zakut, H.		
TITLE	Molecular cloning and construction of the coding region for human acetylcholinesterase reveals a G + C-rich attenuating structure		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.		
MEDLINE	91088577		
REFERENCE	2 (sites)		
AUTHORS	Li, Y., Camp, S., Rachinsky, T.L., Getman, D. and Taylor, P.		
TITLE	Gene structure of mammalian acetylcholinesterase. Alternative exons dictate tissue-specific expression		
JOURNAL	J. Biol. Chem.		
MEDLINE	92078174		
REFERENCE	3 (bases 1 to 1355)		
AUTHORS	Bartels, C.F., Zelinski, T. and Lockridge, O.		
TITLE	Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene accounts for YT blood group polymorphism		
JOURNAL	Am. J. Hum. Genet.		
MEDLINE	93256075		
COMMENT	On Jun 8, 1995 this sequence version replaced gi:348159.		
FEATURES			
SOURCE	**Reference [1] reports bases 1-1224;		
	Reference [J. Biol. Chem. 266, 23083-23090 (1991)] reports bases 1225-1233;		
	Location/Qualifiers		
	1..1355		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="7q22"		
	/tissue_type="blood"		
	<1..136		
	/gene="ACHE"		
	/note="G00-118-746"		
	/number=1		
	137..1224		
	/gene="ACHE"		
	/note="G00-118-746"		
	/number=2		
	1213		
	/gene="ACHE"		
	/note="C [His(YT1)] to A [Asn(YT2)], YT blood group polymorphism"		
	/replace="a"		
VARIATION			
BASE COUNT	207 a 462 c 416 g 270 t		
ORIGIN			

Query Match		62.0%;	Score 1070;	DB 9;	Length 1355;
Local Similarity		100.0%;	Pred. No. 0;		
Matches 1070;		Conservative	0;	Mismatches	0;
QY	1	atgagcccccgaatgctctgtctgacaaagccctccctgctctcccaatctctctc	60		
Db	157	ATGAGCCCCCGCAATGTCCTGCTGCACAGCCTTCCCTGGCTTCCCACTCTTCTCTC	216		
QY	61	ctctctggtctctggtgttgagagagtgagggtcgtgagggcccgagagatgcaagctgtg	120		
Db	217	CTCCCTCTGGCTCTCTGGGTGAGAGATGGGGCGCTGAGAGGCCGGAGAGATGACAGCTGCTG	276		
QY	121	gtgaagagtgctgtgggggcccgtctgggggcatctgcgtgaagaaccccgggggcccctgtc	180		
Db	277	GTGAGGGTGTGGGGGGGGCGCTGGGGGGCAATTCCTGTGAACACCCCGGGGGCCCTGTGC	336		
QY	181	tctgtcttccttggtgatccctcttgccggagcccaatggaaccccgctgctctgtgcca	240		
Db	337	TCTGCTTTCCTGGGATCCCTTTTCGGAGCCACCAATGGAGCCCGCTGCTTTCTGGCA	396		
QY	241	ccggagcccaagcaagcctctgtctcaagggtgtgtagaagctacaaacctccaaagtgtctgc	300		
Db	397	CCGGAGCCCAAGCAGCCTTGGTCAGGGGGGTGAAGCGCTACAACCTTCCAGATGTGCTGC	456		
QY	301	taccaatatgtgaaacccatacccaagttttgagggcaccggagatgtgtgaaccccacaac	360		
Db	457	TACCAATATGTGTGACACCCCTATACCAGGTTTGTGAGGGCACCGACATGTGTGAACCCCAAC	516		
QY	361	cgtagcgtgagaggaagactgtcgtctgacctcaacaggtgtgtgaaacatacccccggcctaca	420		
Db	517	CGTAGCTGTAGCGAGGACTGCTGTACTCAACGTTGTGACACCAATACCCCGGCTTACA	576		
QY	421	tcccccaccctctgctcgtctgtgatactataggggttgagcttcaagtgaggcctctcc	480		
Db	577	TCCCCACCCCGTGCCTGTGTGATCATATGGGGGGGTCTTACAGTGGGGCCCTCTCC	636		
QY	481	ttggaagctgtacgaatgagccgctctctgtgtacagccgagagaactgtctgtgtgtcatg	540		
Db	637	TTTGGACGTGTACGATGAGGGCGGCTTCTTGTGTACAGGCGGAGACGACACTGCTGTCTTCAATG	696		
QY	541	aactccgggtgaggagccttggtctctctgcccctcgccggaggagccgagagcccccgggc	600		
Db	697	AACCTACCGGTGTGGAGGCTTGTGGCTTCCCTGCGCCCTGCGGGAGCCGAAGGCCCCCGGGCT	756		
QY	601	aatgtgggtctctctgatacaagaaggtctggccctcgcaagtgtgtcaagaagaacgttgagacc	660		
Db	757	AATGTGGGTCTCTTGATCTACAGAGCGTGGCCCTCGCATGTGGGTCAAGAGAACGTGGCAGCC	816		
QY	661	ttcgggggtgagccgacatcaagtgaagctgttttgggaagacgagcgagacgcctctgtg	720		
Db	817	TTCCGGGGGTGACCCCGACATACACTGACGCTGTGTGGGAGAGCGGAGACCCGCTCTCGGTG	876		
QY	721	ggcattgacacctgtgttcccccggcccaagccggggccctgttccaaagggtcgctgtgcaagc	780		
Db	877	GGCATTGACACTGTGTCTCCCGCCAGCCGGGGGCGCTGTTCACAGGGCCCTGTGTGCAGAC	936		
QY	781	ggtgcccccaatggaacccctgtggcgcaaggttggtgcatgtggagagggccgtctgaggggcagc	840		
Db	937	GGTGGCCCCCAATGAGACCTCGTGGGCGACAGGTGGGCAATGGAGAGGCCCGTCTGCAGGGCCAGC	996		
QY	841	caagctgagcccaactgttggtgtctctctccaaagcgagcatgtgttggaatgaaacaaagactg	900		
Db	997	CAGCTGGGCCCACTTGTGTGGGCTGTCTCTCCAGCGGGCAGCTGTGTGGGAATGACACAGAGCTG	1056		
QY	901	gttagctgtccttctggaaacagcaacgagcgaggtctccgtgtgaacaaagaaatgagcaagtgtc	960		
Db	1057	GTAGCTCTCTCTTGGAGACAGACCAACGCAAGGTCTGTGTGAATCAACAGGAATGAGCACTGTGCTG	1116		
QY	961	cctcaagaagaagcgtcttccggtctctcctctgtgctgtgtgtgtatgtgaggaactctcagt	1020		
Db	1117	CCTCAAGAAAGAGCTTCTCCGGTCTCTTCTTCTGTGGCTGTGTGTGTGAATGTGAAGACTTCTCTCACT	1176		
QY	1021	gacaccccaagagccctcatcaacgctgggaagacttccacaggtctgcaagt 1070			

Db 1177 GACACCCAGAGGCGCTCATCAACGCGGAGACTTCCACGCGCTGCAGCT 1226

```
RESULT 9
HUMACHEB HUMACHEB 4185 bp DNA linear PRI 23-JUN-1995
LOCUS Homo sapiens acetylcholinesterase (ACHE) gene, exons 2-6.
DEFINITION L42812
ACCESSION L42812.1 GI:854682
VERSION acetylcholinesterase.
KEYWORDS Homo sapiens DNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Bartels,C.F., Zelinski,T. and Lockridge,O.
TITLE Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene
accounts for YF blood group polymorphism
JOURNAL Am. J. Hum. Genet. 52 (5): 928-936 (1993)
MEDLINE 93256075
REFERENCE 2 (bases 1 to 4185)
AUTHORS Bartels,C.F., Moriarty,P.L., Becker,R.E., Robbs,R.S.,
Sorenson,R.C., Montjoy,C.P. and Lockridge,O.
TITLE Polymorphic sites in the acetylcholinesterase gene of patients with
Alzheimer's disease
JOURNAL Unpublished (1995)
FEATURES
    source      1..4185
                Location/Qualifiers
                1..4185
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /map="7q22"
                /cell_type="white blood cell"
                /gene="ACHE"
                /note="G00-118-746"
                /number=2
                /evidence=experimental
    gene        75..4185
                /evidence=experimental
                /gene="ACHE"
                /number=1
    variation   1151
                /gene="ACHE"
                /note="YF blood group; His322/Asn"
                /replace="a"
                /number=1
    intron      1163..1508
                /gene="ACHE"
                /note="G00-118-746"
                /number=2
                /evidence=experimental
    exon        1509..1993
                /gene="ACHE"
                /note="G00-118-746"
                /number=3
                /evidence=experimental
    variation   1871
                /gene="ACHE"
                /note="Pro446/Pro"
                /replace="t"
                /number=1
    intron      1994..2988
                /gene="ACHE"
                /note="G00-118-746"
                /number=3
                /evidence=experimental
    variation   2309
                /gene="ACHE"
                /note="g or a; within Alu repeat; g more common;
                /polymorphism at 2309"
                /replace="a"
                /number=1
    exon        2989..3158
                /gene="ACHE"
                /note="G00-118-746"
                /number=4
                /evidence=experimental
```

```
intron      3159..3238
                /gene="ACHE"
                /note="G00-118-746"
                /number=4
                /evidence=experimental
    exon      3239..3991
                /gene="ACHE"
                /note="Includes retained intron 5; G00-118-746"
                /number=5
                /evidence=experimental
    variation 3290
                /gene="ACHE"
                /note="Pro 561 more common; C/G polymorphism at 3290;
                /Pro561(CCG) or Arg 561(CGG)"
                /replace="g"
                /number=1
    exon      3992..4185
                /gene="ACHE"
                /note="G00-118-746"
                /number=6
                /evidence=experimental

BASE COUNT  792 a 1320 c 1236 g 837 t
ORIGIN
Query Match 62.0%; Score 1070; DB 9; Length 4185;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagccccagagtgctgctgcaagacgttcctcgtgctccacgtcttcctc 60
Db 95 atgagccccagagtgctgctgcaagacgttcctcgtgctccacgtcttcctc 154
QY 61 ctctctgctccttgaggtagagagtgaggagctgaggccgaggagatgcagagctgctg 120
Db 155 ctctctgctccttgaggtagagagtgaggagctgaggccgaggagatgcagagctgctg 214
QY 121 gtgacggtgcgtgaggcgagctgcggggcattcgctgtagaagaccccgaggcctgttc 180
Db 215 gtgacggtgcgtgaggcgagctgcggggcattcgctgtagaagaccccgaggcctgttc 274
QY 181 tctgcttccttgaggatccctctgcggagaccccaatggagaccggtgcttcctgcga 240
Db 275 tctgcttccttgaggatccctctgcggagaccccaatggagaccggtgcttcctgcga 334
QY 241 ccggagcccaagagccttgctgaggggtgtagacgtacaaacctccagagtgcttcgc 300
Db 335 ccggagcccaagagccttgctgaggggtgtagacgtacaaacctccagagtgcttcgc 394
QY 301 taccataatgtgaacacctatacccaaggtcttgaggagccgagatgtgaaacccaac 360
Db 395 taccataatgtgaacacctatacccaaggtcttgaggagccgagatgtgaaacccaac 454
QY 361 cgtgagctgagcgagagctgcctgtacctcaacggtgtagacacatacccccgagctaca 420
Db 455 cgtgagctgagcgagagctgcctgtacctcaacggtgtagacacatacccccgagctaca 514
QY 421 tccccaccctgttcctcgctcgtgacatcaggggtggtctctacagtgaggcctctcc 480
Db 515 tccccaccctgttcctcgctcgtgacatcaggggtggtctctacagtgaggcctctcc 574
QY 481 ttgagagtgtagagatgagcgtcctctgtgtagcagggcgagagagctgtgtgtgtcatg 540
Db 575 ttgagagtgtagagatgagcgtcctctgtgtagcagggcgagagagctgtgtgtgtcatg 634
QY 541 aactaccgggtgaggacgttgcttcctcgtgacccctgcggaggagccgagagcccgagg 600
Db 635 aactaccgggtgaggacgttgcttcctcgtgacccctgcggaggagccgagagcccgagg 654
QY 601 aatgtggtctccttgagatcagagagctgagcctgcaagtgggtgaggaagcgtgagcagc 660
Db 695 aatgtggtctccttgagatcagagagctgagcctgcaagtgggtgaggaagcgtgagcagc 754
QY 661 ttcgggggtgagccgagacatagtgacgctgtttggggagagcgcgaggagccgctcgtg 720
```

```

Db 755 TTCCGGGGTGCACCCGACATCAGTGCCTGTTGGGGAGACGGGGAGCCGCTCGGTG 814
Oy 721 ggcacgtcactctgtctcccccagccgggacctgtctccacagggccgtgtgcaagc 780
Db 815 GGCATGACACCTGCTGCTCCGCCGACGCGGGGCGCTTCACAGGGCGGCTGCTGCAGAGC 874
Oy 781 ggtgtcccccacatgacacctgtgacagagtgatgagagagccgtcgcagagggcagc 840
Db 875 GGTGCCCCCAATGAGACCTGTGGCCACGCTGGCATGGAGAGGCCCTGTGCAGAGGCCACG 934
Oy 841 cagctggcccccactgtgtggtgtctctccacagggacctgtgtggaatgacacagagtg 900
Db 935 CAGCTGGCCCCACCTGTGTGGGCTGTCTCCACGCGGACTGTGGGAATGACACAGAGCTG 994
Oy 901 gtaagctcctctggacacagcagcagcagagtgctctgtgtaacacagaaatggcacgtgtg 960
Db 995 GTACCTGCTGCTTCGGACACGACGACGAGCTGCTGCTGTCGTAACGAGATGGCAGCTGCTG 1054
Oy 961 cctcaagaagcgtctctcgtctctctctctctctctgtgtgtagatgagacctctcagc 1020
Db 1055 CCTCAGAGAACGCTCTTCGGGCTCTCTCTCTCTCTGCTGTGTAGATGAGACTTCTCTCAGT 1114
Oy 1021 gacacccacagagccctcatcaacggcgagagactccagcgcctgcagagt 1070
Db 1115 GACACCCACGAGGCCCTCATCAGCGGGGACACTTCCAGGCGCTGCAGGT 1164

RESULT 10
AC011895/c 172358 bp DNA linear PRI 30-SEP-2000
LOCUS AC011895
DEFINITION Homo sapiens BAC clone RP11-126L15 from 7q22, complete sequence.
ACCESSION AC011895
VERSION AC011895.4 GI:9454624
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulhelia; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Sulston,J.E. and Waterston,R.
TITLE The sequence of Homo sapiens BAC clone RP11-126L15
JOURNAL Unpublished
MEDLINE 3 (bases 1 to 172358)
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
-----
Genome Center

```

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0126L15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, send <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateo, M., Catanesi, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pterier de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-336D7. Actual start of this clone is at base position 1 of RP11-126L15; actual end is at base position 172358 of RP11-126L15.

The clone RP11-126L15 contains a transposon that inserts at base position 30153, which has been omitted from the submitted sequence.

There are polymorphic base pair differences in the overlap between the clone RP11-126L15 and RP11-336D7.
 Location/Qualifiers

```

FEATURES
  source
    1..172358
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="7q22"
    /clone="RP11-126L15"
    /clone_1lb="RPCI-11"
    repeat_region
      1..249
      /rpt_family="Alu"
    repeat_region
      935..1075
      /rpt_family="Alu"
    repeat_region
      1076..1390
      /rpt_family="Alu"
    repeat_region
      1391..1553
      /rpt_family="Alu"
    repeat_region
      2087..2371
      /rpt_family="Alu"
    repeat_region
      3027..3324
      /rpt_family="Alu"
    repeat_region
      3357..3491
      /rpt_family="Alu"

```


|||||
Db 135524 CAGCTGGCCCACTTGTGGCTGCTCCAGCGGCACTGGTGGATGACACAGAGCTG 135465
Oy 901 gtagcctgcttcggacacgacgacgagcaggttcctgtgtgacacgaaatgagcagctgtg 960
Db 135464 GTAGCCTGCTTCGACACGACGACGAGGTCCTGTGTAACACAGATGACAGCTGCTG 135405
Oy 961 cctcaagaagcgtcttcctgcttcctgctgtgctgtgtgtagatgagagactcctcagt 1020
Db 135404 CCTCAAGAAACGCTTCCTCGGCTTCCTGCTGCTGTGTGATGAGAGACTTCCTCAGT 135345
Oy 1021 gacacccagagagcctcatcaagcgaggagactccacgagcctcaggt 1070
Db 135344 GACACCCAGAGGCGCTCATCAACGCGGAGACTTCACAGGCGCTCAGGT 135295
RESULT 11
AC084057
LOCUS AC084057 194681 bp DNA linear HTG 12-OCT-2000
DEFINITION Homo sapiens chromosome 7 clone CTC-786K12, WORKING DRAFT SEQUENCE,
36 unordered pieces.
AC084057
AC084057.1 GI:10799406
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 194681)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 194681)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 2351286
Center clone name: CTC-786K12

Summary Statistics
Consensus quality: 162650 bases at least Q40
Consensus quality: 173899 bases at least Q30
Consensus quality: 176608 bases at least Q20
Estimated insert size: 200000; pulse field gel estimation
Estimated insert size: 191181; sum-of-contigs estimation
Quality coverage: 7.95 in Q20 bases; pulse field gel estimation
Quality coverage: 8.32 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
consists of 36 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 1324: contig of 1324 bp in length
1325 1424: gap of unknown length
1425 2586: contig of 1162 bp in length
2587 2687: gap of unknown length
2688 4056: contig of 1370 bp in length
4057 4157: gap of unknown length
4158 5197: contig of 1041 bp in length
5198 5298: gap of unknown length
5299 6414: contig of 1116 bp in length
6415 7675: gap of unknown length
7676 7776: contig of 1068 bp in length
7777 8844: gap of unknown length
8845 9998: contig of 1055 bp in length
9999 10099: gap of unknown length
10100 11598: contig of 1500 bp in length
11599 11699: gap of unknown length
11700 12957: contig of 1258 bp in length
12958 13057: gap of unknown length
13058 14169: contig of 1112 bp in length
14170 14269: gap of unknown length
14270 15288: contig of 1019 bp in length
15289 15387: gap of unknown length
15388 16937: contig of 1549 bp in length
16938 17037: gap of unknown length
17038 18087: contig of 1050 bp in length
18088 18187: gap of unknown length
18188 19428: contig of 1241 bp in length
19429 19528: gap of unknown length
19529 20571: contig of 1043 bp in length
20572 20671: gap of unknown length
20672 22194: contig of 1523 bp in length
22195 22294: gap of unknown length
22295 24474: contig of 2180 bp in length
24475 24574: gap of unknown length
24575 25996: contig of 1422 bp in length
25997 26096: gap of unknown length
26097 27844: contig of 1748 bp in length
27845 27944: gap of unknown length
27945 30121: contig of 2177 bp in length
30122 30221: gap of unknown length
30222 31961: contig of 1741 bp in length
31962 32061: gap of unknown length
32062 35719: contig of 3658 bp in length
35720 35819: gap of unknown length
35820 38072: contig of 2252 bp in length
38073 38171: gap of unknown length
38172 41171: contig of 3000 bp in length
41172 41271: gap of unknown length
41272 47552: contig of 6281 bp in length
47553 47653: gap of unknown length
47654 52017: contig of 4365 bp in length
52018 52117: gap of unknown length
52118 55609: contig of 3492 bp in length
55610 55710: gap of unknown length
55711 60280: contig of 4570 bp in length
60281 60379: gap of unknown length
60380 66246: contig of 5867 bp in length
66247 66347: gap of unknown length
66348 78184: contig of 11838 bp in length
78185 78285: gap of unknown length
78286 96185: contig of 17900 bp in length
96186 96285: gap of unknown length
96286 115729: contig of 19445 bp in length
115730 115829: gap of unknown length
115830 139049: contig of 23220 bp in length
139050 139149: gap of unknown length
139150 161902: contig of 22753 bp in length
161903 162002: gap of unknown length
162003 194681: contig of 32679 bp in length.
Location/Qualifiers
1. 194681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/clone="CTC-786K12"
/clone_11b="CalTech human BAC library C"
ORIGIN 45257 a 49777 c 50799 g 45331 t 3517 others

Query Match 62.0%; Score 1070; DB 2; Length 194681;
Best Local Similarity 100.0%; Pred. No. 0;

RESULT	14
LOCUS	HUMACHE03 510 bp DNA linear PRI 14-JUN-1995
DEFINITION	Human acetylcholinesterase (ACHE) gene, exons 4 and 5.
ACCESSION	L22561
VERSION	L22561.1 GI:862301
KEYWORDS	acetylcholinesterase.
SEGMENT	3 of 4
SOURCE	Homo sapiens DNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (sites)
TITLE	Soreq,H.E., Ben-Aziz,R., Prody,C.A., Seldman,S., Gnatt,A., Neville,L., Lieman-Hurwitz,J., Lev-Lehman,E., Ginzberg,D., Lapidot-Lifson,Y. and Zakut,H. Molecular cloning and construction of the coding region for human acetylcholinesterase reveals a G + C-rich attenuating structure Proc Natl. Acad. Sci. U.S.A. 87 (24), 9688-9692 (1990)
JOURNAL MEDLINE	91088577
REFERENCE	2 (sites)
AUTHORS	Li,Y., Camp,S., Rachinsky,T.L., Gelman,D. and Taylor,P.
TITLE	Gene structure of mammalian acetylcholinesterase. Alternative exons dictate tissue-specific expression J. Biol. Chem. 266 (34), 23083-23090 (1991)
JOURNAL MEDLINE	92078174
REFERENCE	3 (bases 1 to 510)
AUTHORS	Bartels,C.F., Zelinski,T. and Lockridge,O.
TITLE	Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene accounts for XT blood group polymorphism Am. J. Hum. Genet. 52 (5), 928-936 (1993) 9356075
JOURNAL MEDLINE	
COMMENT	On Jun 15, 1995 this sequence version replaced gi:853591. **Reference [1] reports bases 106-275; Reference [J. Biol. Chem. 266, 23083-23090 (1991)] reports bases 276-507; Reference [3] reports bases 1-105**.
FEATURES	Location/Qualifiers
SOURCE	1..510
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/map="7q22"
	/tissue_type="blood"
Intron	order(L22560.1:583..>779,<1..106) /note="G00-118-746"
	/number=3
exon	107..276 /gene="ACHE"
	/note="G00-118-746"
	/number=4
Intron	277..356 /gene="ACHE"
	/note="G00-118-746"
	/number=4
exon	357..>510 /gene="ACHE"
	/note="glycollipid anchored form; G00-118-746"
	/number=5
variation	408 /gene="ACHE"
	/note="C to G (Pro561 to Arg) polymorphism"
	/replace="g"
BASE COUNT	65 a 204 c 150 g 91 t
ORIGIN	
Query Match	9.9%; Score 170; DB 9; Length 510;
Best Local Similarity	100.0%; Pred. No. 1,4e-81;
Matches 170; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1553 ggagatcccaatgagcggcgagccccaaaggccccacaatgcccgcctacacgcgcgagg 1612
DB	106 ggagatcccaatgagcggcgagccccaaaggccccacaatgcccgcctacacgcgcgagg 165

Oy	1613	ctgaagcagtacgtatgactctgtagacctcgsggcgcgttggaaagtgtccggcggggcgccgcgcgc	1672
Db	166	CTCAGCAAGTACGTTAAGTCGTGACCCTGGCGCCGTGGAGAGTGCGCGGGGCGGCCTGCCGCC	225
Oy	1673	aggcctgcgcctcttggaaccgccttcctccccaattgctcaagcgccacc	1722
Db	226	AGGCCTGCGCCTTCTGTGGAAACCGCTTCTCCCAAATTGCTCAAGCGCCACC	275
RESULT 15			
AFJ12032			
LOCUS			
DEFINITION	AFJ12032	117962 bp DNA linear PRI 20-MAR-2001	
	Homo sapiens acetylcholinesterase (ACHE) gene, partial cds; ASR2		
	(ASR2) gene, complete cds, alternatively spliced; thyrold receptor		
	interacting protein 6 (TRIP6) gene, complete cds; cation-chloride		
	cotransporter (CICP1) gene, complete cds, alternatively spliced;		
	ephrin type-B receptor 4 precursor (EPHB4) and zonadhesin (ZAN)		
	genes, complete cds; and unknown gene.		
ACCESSION	AFJ12032		
VERSION	AFJ12032.1	GI:13383497	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Euthioria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 117962)		
AUTHORS	Wilson,M.D., Riemer,C., Martindale,D.W., Schnupf,P., Boright,A.P.,		
	Cheung,T.L., Hardy,B.M., Schwartz,S., Scherer,S.W., Tsui,L.-C.,		
	Miller,W. and Koop,B.F.		
TITLE	Comparative analysis of the gene-dense ACHE/TRF2 region on human		
JOURNAL	Chromosome 7q22 with the orthologous region on mouse chromosome 5		
MEDLINE	Nucleic Acids Res. 29 (6), 1352-1365 (2001)		
REFERENCE	2 (bases 1 to 117962)		
AUTHORS	Wilson,M.D., Martindale,D.W. and Koop,B.F.		
TITLE	Direct Submission		
JOURNLT	Submitted (06-OCT-2000) Biology, Centre for Environmental Health,		
	University of Victoria, P.O. Box 3020, Victoria, B.C. V8W 3M5,		
	Canada		
FEATURES			
source	Location/Qualifiers		
	1..117962		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="7"		
	/map="7q22"		
	/clone="cos159d9.T and H.DJ013bm12 fusion"		
	join(<1..165,999..1339)		
	/gene="ACHE"		
	/product="acetylcholinesterase"		
	<1..1339		
gene	/gene="ACHE"		
	join(<1..165,999..1120)		
CDS	/gene="ACHE"		
	/codon_start=3		
	/product="acetylcholinesterase"		
	/protein_id="AAK21003.1"		
	/db_xref="GI:13383498"		
	/translation="NEPRPKAPQWPPTTAGAQQVYSLDRLPLEVRGRGLRAOACAFNN		
	RFLPKLSATDTLDEAEIRWKAEFRWSSYVMWNQPDHYSKDRCSDL"		
	1676..2594		
MRNA	/product="unknown"		
	2048..2476		
CDS	/codon_start=1		
	/evidence=no experimental		
	/product="unknown"		
	/protein_id="AAK21004.1"		
	/db_xref="GI:13383499"		
	/translation="NGDKRPFGRSRDNIIGCYEASISLCIAHFGEQGQRGLCHPVGVGH		
	ELGAERLISHFAGGCVFMVGADARSALLGVCVSGSTENVVLDPDHVWGTPKSPS		
	ETDAQAWGMQEVSAAFDNSFNLCITLS5000ORLD"		
	complement(join(12655..2845,2936..3062,3176..3265,		
	3460..3616,3806..3999,4119..4272,4351..4524,4884..5072		
MRNA			

5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15746,15990. .16352))
/gene="ASR2"
/note="alternatively spliced; missing 12 bp from exon 18"
/product="ASR2B"
complement(join(2655. .2845,2936. .3062,3176. .3377,
3460. .3616,3806. .3999,4119. .4272,4351. .4524,4884. .5072,
5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15746,15990. .16352))
/gene="ASR2"
/note="alternatively spliced"
/product="ASR2A"
complement(2655. .16352)
/gene="ASR2"
complement(join(2770. .2845,2936. .3062,3176. .3277,
3460. .3616,3806. .3999,4119. .4272,4351. .4524,4884. .5072,
5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15728))
/gene="ASR2"
/note="alternatively spliced"
/codon_start=1
/protein_id="AAK21005.1"
/db_xref="GI:13383500"
/translation="MGSDSDDEYDRRRDRFRERSDYDRSREHDERRRGGDNDNRMD
RGRERSGEYRDYDRNRREDFSPRHLSPPQKRRMDDEHSDPYSGYMPAG
GGGPTGPPQPGHDPVIMOHVLPFLARGLSIDGLGVPPEYMKFKKRLSLD
DSVDEAVKRYNDYKLDPRRODMODEFLAHRDEMFRRSKYHPDEYKROEAKLO
NRLRVLSLMEGTGFNLLDIDKADAIYKMDAIVKMGTEENDRLTIDEESEQ
ACGPGPSKKEBGRAGAGLDGERTKNDDEKEDQAEENDSNDKTKKSGDGD
EKKEDSEKAKSKRNKRKSGDSDPBGSESESESGOAEKEBEAEALKE
KEKPEEWEKPKDAGLECKPRPLKTKSLFMRNIAPNISRAEIIISLKRYFARN
ALSEPQERREFRGWTFDRSVNIKEICMONIRLRECELSPGVNRDTRVRIN
GITOHKOIYRNDIKLAALKLHLTDORTLOLMASEPPPLPTSLPSONPLIKNTDILI
EYSAEEELSSSGAGPPEPEKGNPAINEVERDEKLKYLIDKLILYRTHSDY
YNTCEYFNDEMPNRCGIIHVRGMPNPNISHEVLEWQTEFEKLPPLISVLSSE
EEAKQKGRDDEQVEKFTVNTQELGDKMCLPSGKKFKGPEFYRKHIFNKHAKI
EYKKEVAFNNELTDAKRPALPEIKPAQPGCAQILPGLTFLPYHOTQGLMRY
GPRRPILGAGAVRAPVTGTPPYPHAPYGAGRCNYDAFRQGGYPGKPRRMVRG
DPAIYERDLADPDVDFE"
complement(join(2770. .2845,2936. .3062,3176. .3365,
3460. .3616,3806. .3999,4119. .4272,4351. .4524,4884. .5072,
5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15728))
/gene="ASR2"
/note="alternatively spliced; missing 12 bp from exon 18"
/codon_start=1
/protein_id="AAK21006.1"
/db_xref="GI:13383501"
/translation="MGSDSDDEYDRRRDRFRERSDYDRSREHDERRRGGDNDNRMD
RGRERSGEYRDYDRNRREDFSPRHLSPPQKRRMDDEHSDPYSGYMPAG
GGGPTGPPQPGHDPVIMOHVLPFLARGLSIDGLGVPPEYMKFKKRLSLD
DSVDEAVKRYNDYKLDPRRODMODEFLAHRDEMFRRSKYHPDEYKROEAKLO
NRLRVLSLMEGTGFNLLDIDKADAIYKMDAIVKMGTEENDRLTIDEESEQ
ACGPGPSKKEBGRAGAGLDGERTKNDDEKEDQAEENDSNDKTKKSGDGD
EKKEDSEKAKSKRNKRKSGDSDPBGSESESESGOAEKEBEAEALKE
KEKPEEWEKPKDAGLECKPRPLKTKSLFMRNIAPNISRAEIIISLKRYFARN
ALSEPQERREFRGWTFDRSVNIKEICMONIRLRECELSPGVNRDTRVRIN
GITOHKOIYRNDIKLAALKLHLTDORTLOLMASEPPPLPTSLPSONPLIKNTDILI
EYSAEEELSSSGAGPPEPEKGNPAINEVERDEKLKYLIDKLILYRTHSDY
YNTCEYFNDEMPNRCGIIHVRGMPNPNISHEVLEWQTEFEKLPPLISVLSSE
EEAKQKGRDDEQVEKFTVNTQELGDKMCLPSGKKFKGPEFYRKHIFNKHAKI
EYKKEVAFNNELTDAKRPALPEIKPAQPGCAQILPGLTFLPYHOTQGLMRY
GPRRPILGAGAVRAPVTGTPPYPHAPYGAGRCNYDAFRQGGYPGKPRRMVRG
DPAIYERDLADPDVDFE"
complement(join(17865. .18146,18574. .18694,19596. .19774,
20574. .20743,20854. .20947,22453. .22824,23086. .23211,
23331. .23458,23713. .23970))

gene
/gene="TRIP6"
/product="thyroid receptor interacting protein 6"
complement(17865. .23970)
/gene="TRIP6"
complement(join(18015. .18146,18574. .18694,19596. .19774,
20574. .20743,20854. .20947,22453. .22824,23086. .23211,
23331. .23458,23713. .23821))
/gene="TRIP6"
/codon_start=1
/product="thyroid receptor interacting protein 6"
/protein_id="AAK21007.1"
/db_xref="GI:13383502"
/translation="MSCPTMLPPKQEPAPAPGRIAPCTGPPAPAHANAPHPRV
NFCPLSBOCTOAPGPPEDRGAPVASHGVLDTQGLPADRGSLRFSGLDAEIDLSS
TLAELNGGCHASRRPDRAPVPPAPRTGSLPKNPASPLPASPAGPTPASYTTA
STPAGAPFPOVQVAPVRCGCPRRGASASQPLRPHPLRPGREVGPGYRSORE
PCGAPAEAGAVSGPAGRCGCGHPVPLSGPDEBDRITKLVNDMNPSPGCVF
GCGGGGGEVDVGGAGVALLDRVFNHGCFTSCRAOLGQNHVATERRAVGCVVA
TLEKCKASQPTLIDRLPAMKRYHPGCGCTCYCHRGIDGIPPTVATSOIKIEDFH
RKAPRCVGGAGIMEPGOETVRIVALDRSFHICVCEBGLLSBGEQCTP
LDGHLCKACSAWRIQELSAVTTDC"
complement(join(24306. .25600,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38581))
/gene="CIP1"
/note="alternatively spliced"
/product="cation-chloride cotransporter-interacting
protein 1"
complement(24306. .38581)
/gene="CIP1"
complement(join(24714. .25600,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/codon_start=1
/product="cation-chloride cotransporter-interacting
protein 1"
/protein_id="AAK21008.1"
/db_xref="GI:13383503"
/translation="MASESPLIAYLLIGEVALPAGGAGGASARKISTELIGV
VPIVLSMFSIVPFLRGVSGHGLALAMKILAFILATVYLSGATATNAGVG
GATMTSRLTGPVSGSLGIMFLTAVNCCASLVLSEVLDVFEADTGBGLVL
POQYGNLLYGSLLDVGVCYTLGAGLARAASPLTVSGSLVLSLFAVAGRD
IRLTTPPFGNSLPPRGFTGFTGNSGTLDNGAGYAEYTTGAVMNASVAVLEN
GCTGIMAGNMSELKDPRAIPLGTIVAAVYFFVYVLLFSLSTFCRTLOEDYG
FPRALSLMPPLVIGITATLASMSLSLIGASLILHALRDLEFYILAPKVVSRG
NPAALVYSNGVQVILAGIKNTLAAYVTFVAVYAVDLSCLSEMASPNRPPT
FELFSMHTCLGVASCLLAMEFLISPRAGSGLLKLALALTARGPSWGVGAL
LEHQVKYLLRLDVRDHYKFWRPQLLLLVGNRGLPLRLANQHLKGLVYLGHT
LGLDLPLSDPVQPOYGAMLSIDRAOVAFVDTLPSVROAQLRLISIGLGMK
NLTVLGFDYDAPQDHFLLDPAFSEPDSTRESSPLSTLPPRAPSPRALNPD
YVATVADLKKMNTVILARSGALPPEPRLSGSGGSHHDWALNLTLRGPGGY
VIVCGFLQMKNTTIGMVPWASHARIRLTCISGRAPRAAEGRALWALISQRIARV
QEVYREGGAGPEPAEEDGFVNSRGAEBALAKSNALYRAQCGCTGGPGCP
EGGDAGCPITALTFLVLPRPADPARYPRILALETLTRDLGPTLLVHVOTPTCDL
"
complement(join(25787. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))

Query Match 9.5%; Score 164; DB 9; Length 117962;
Best Local Similarity 100.0%; Pred. No. 2.5e-78;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1559 ccaatgagcccgagaccacaaatgagcccgatcacggcgggggtcgaac 1618
|||||

Db	1	CCATGAGCCCCGGAGACCCCAAGGCCCAAAATGGCCCCCGGTCACACGGCGGGGGCTCAGC	60
Qy	1619	AGTGTCTAGTCTGGAGACTGGGCGCGCTGAGAGTGGGAGGGAGCTGTGGCCCCAGAGCT	1678
Db	61	AGTACGTGTAAGTGGAGACTCGGGCCCCCTGGAGAGTGGCGGGGGGCTGCGCCCCAGAGCTT	120
Qy	1679	GGAGCTCTGGAGACGCTCTCTCTCCCAAAATGTCTCAGCGCAAC	1722
Db	121	GGCGCTCTTGGAGACGCTCTCTCTCTCCCAAAATGTCTCAGCGCCACC	164

Search completed: August 31, 2002, 22:49:43
Job time: 26012 sec

THIS PAGE BLANK (USPTO)

BASE COUNT 5 a modified to introduce an Nco I restriction site
 ORIGIN 11 c 8 g 5 t

Query Match 100.0%; Score 29; DB 6; Length 29;
 Best Local Similarity 100.0%; Pred. No. 0.0041;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatattcagcagcagctagagcccccgc 29
 |||||||
 Db 1 GATATTCAGCAGCAGCTAGAGCCCCC 29

RESULT 2
 LMFCHR36_26
 WPCOMMENT

Sequence split into 36 fragments Locus LMFCHR36 Accession AL499624

Fragment Name	Begin	End
LMFCHR36_00	1	110000
LMFCHR36_01	100001	210000
LMFCHR36_02	200001	310000
LMFCHR36_03	300001	410000
LMFCHR36_04	400001	510000
LMFCHR36_05	500001	610000
LMFCHR36_06	600001	710000
LMFCHR36_07	700001	810000
LMFCHR36_08	800001	910000
LMFCHR36_09	900001	1010000
LMFCHR36_10	1000001	1110000
LMFCHR36_11	1100001	1210000
LMFCHR36_12	1200001	1310000
LMFCHR36_13	1300001	1410000
LMFCHR36_14	1400001	1510000
LMFCHR36_15	1500001	1610000
LMFCHR36_16	1600001	1710000
LMFCHR36_17	1700001	1810000
LMFCHR36_18	1800001	1910000
LMFCHR36_19	1900001	2010000
LMFCHR36_20	2000001	2110000
LMFCHR36_21	2100001	2210000
LMFCHR36_22	2200001	2310000
LMFCHR36_23	2300001	2410000
LMFCHR36_24	2400001	2510000
LMFCHR36_25	2500001	2610000
LMFCHR36_26	2600001	2710000
LMFCHR36_27	2700001	2810000
LMFCHR36_28	2800001	2910000
LMFCHR36_29	2900001	3010000
LMFCHR36_30	3000001	3110000
LMFCHR36_31	3100001	3210000
LMFCHR36_32	3200001	3310000
LMFCHR36_33	3300001	3410000
LMFCHR36_34	3400001	3510000
LMFCHR36_35	3500001	3529852

Continuation (27 of 36) of LMFCHR36 from base 2600001 (AL499624) Leishmania major chromo

Query Match 72.4%; Score 21; DB 2; Length 110000;
 Best Local Similarity 82.8%; Pred. No. 37;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 gatattcagcagcagctagagcccccgc 29
 |||||||
 Db 78633 GATATTCAGCAGCAGCTAGAGCCCCC 78661

RESULT 3
 AC015957 157813 bp DNA 11near HTG 14-MAR-2000
 LOCUS Homo sapiens chromosome 18 clone RP11-51B9 map 16, WORKING DRAFT
 DEFINITION
 SEQUENCE, 3 unordered pieces.
 ACCESSION AC015957

VERSION AC015957.3 GI:7108397
 KEYWORDS HTG; PHASEL; HTG_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE 1 (bases 1 to 157813)
 JOURNAL Homo sapiens chromosome 18, clone RP11-51B9
 REFERENCE 2 (bases 1 to 157813)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Barina, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeRubeis, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Harford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Lebeck, J., Liu, C., Locke, K., MacDonald, P., Marquis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
 Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Tesfaye, S., Tittell, A., Vassiliou, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Feb 28, 2000 this sequence version replaced gi:1579286.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L959
 Center clone name: 51_B_9

Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 0.047783086478390Chemistry: Dye-primer-amerism; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 138215 bases at least Q40
 Consensus quality: 150140 bases at least Q30
 Consensus quality: 155339 bases at least Q20
 Insert size: 163000; agarose-efp
 Insert size: 157613; sum-of-coverage

Quality cover.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 3 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved
 13053: contig of 13053 bp in length
 13054 13153: gap of 100 bp
 13154 51004: contig of 37851 bp in length
 51005 51104: gap of 100 bp
 51105 157813: contig of 106709 bp in length.

FEATURES
 SOURCE
 1. 157813
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="RP11-51B9"
 /clone_lib="RPCT-11 Human Male BAC"

```

misc_feature      1..13053
                  /note="assembly_fragment"
misc_feature      13154..51004
                  /note="assembly_fragment"
                  /clone_end:SP6
                  /vector_side:right"
                  51105..157813
                  /note="assembly_fragment"
                  /clone_end:T7
                  /vector_side:left"
BASE COUNT      50360 a 31888 c 30054 g 45308 t 203 others
ORIGIN
Query Match      69.7%, Score 20.2, DB 2, Length 157813;
Best Local Similarity 88.0%; Pred. No. 92;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 5 tctgcagccatgctcagggcccccgc 29
      |||||
Db 110234 tctgcagccatgctcagggcccccgc 110258

RESULT 4
AC013700
LOCUS            AC013700
DEFINITION      Homo sapiens clone RP11-22C7, WORKING DRAFT SEQUENCE, 14 unordered
                  pieces.
ACCESSION        AC013700
VERSION          AC013700.4 GI:7259721
KEYWORDS         HTGS, HTGS_PHASE1, HTGS_DRAFT.
SOURCE            human.
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE        1 (bases 1 to 169979)
AUTHORS          Birten,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE            Homo sapiens, clone RP11-22C7
JOURNAL           Unpublished
PUBLISHED        2 (bases 1 to 169979)
AUTHORS          Birten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
                  Baldwin,J., Barna,N., Beckery,R., Boguslavsky,L., Bouknighter,B.,
                  Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
                  Cooke,P., Dextrallano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
                  Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
                  Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hesford,A., Horton,L.,
                  Howell,J.C., Johnson,R., Jones,C., Kann,L., Karats,A., Klein,J.,
                  Lohocky,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,
                  McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
                  Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
                  Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
                  Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
                  Teffaye,S., Tirelli,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X.,
                  Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE            Direct Submission
JOURNAL           Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
                  Research, 320 Charles Street, Cambridge, MA 02141, USA
                  On Mar 17, 2000 this sequence version replaced gi:6970506.
COMMENT          All repeats were identified using RepeatMasker:
                  Smt, A.F.A. & Green, P. (1996-1997)
                  http://ftp.genome.washington.edu/RM/RepeatMasker.html
                  ----- Genome Center
                  Center: Whitehead Institute/ MIT Center for Genome Research
                  Center code: MIBR
                  Web site: http://www-seq.wi.mit.edu
                  Contact: sequence_submissions@genome.wi.mit.edu
                  ----- Project Information
                  Center project name: L4099
                  Center clone name: 22.C7
                  ----- Summary Statistics
                  Sequencing vector: M13; M77815; 100% of reads
                  Chemistry: Dye-terminator Big Dye; 100% of reads
                  Assembly program: Phrap; version 0.960731

```

```

Consensus quality: 160794 bases at least Q40
Consensus quality: 165604 bases at least Q30
Consensus quality: 167325 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 168679; sum-of-contigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1339: contig of 1339 bp in length
* 1340 1439: gap of 100 bp
* 1440 2542: contig of 1103 bp in length
* 2543 2642: gap of 100 bp
* 2643 4104: contig of 1462 bp in length
* 4105 4204: gap of 100 bp
* 4205 5645: contig of 1441 bp in length
* 5646 5745: gap of 100 bp
* 5746 10797: contig of 5052 bp in length
* 10798 10897: gap of 100 bp
* 10898 18365: contig of 7468 bp in length
* 18366 18465: gap of 100 bp
* 18466 28722: contig of 10257 bp in length
* 28723 28822: gap of 100 bp
* 28823 40013: contig of 11191 bp in length
* 40014 40113: gap of 100 bp
* 40114 53470: contig of 13357 bp in length
* 53471 53570: gap of 100 bp
* 53571 69956: contig of 16386 bp in length
* 69957 70056: gap of 100 bp
* 70057 85519: contig of 15463 bp in length
* 85520 85619: gap of 100 bp
* 85620 108587: contig of 22968 bp in length
* 108588 108687: gap of 100 bp
* 108688 134650: contig of 25963 bp in length
* 134651 134750: gap of 100 bp
* 134751 169979: contig of 35229 bp in length.
Location/Qualifiers
FEATURES
source
1..169979
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="RP11-22C7"
/clone_lib="RPC1-11 Human Male BAC"
1..1339
/misc_feature
/note="assembly-fragment"
1440..2542
/misc_feature
/note="assembly-fragment"
2643..4104
/misc_feature
/note="assembly-fragment"
4205..5645
/misc_feature
/note="assembly-fragment"
5746..10797
/misc_feature
/note="assembly-fragment"
10898..18365
/misc_feature
/note="assembly-fragment"
18466..28722
/misc_feature
/note="assembly-fragment"
28823..40013
/misc_feature
/note="assembly-fragment
clone_end:SP6
vector_side:left"
40114..53470
/misc_feature
/note="assembly-fragment"
53571..69956
/misc_feature
/note="assembly-fragment"
70057..85519
/misc_feature
/note="assembly-fragment"

```

```

misc-feature      85620..108587
                  /note="assembly-fragment"
misc-feature      108688..134650
                  /note="assembly-fragment"
misc-feature      134751..169979
                  /note="assembly-fragment"
BASE COUNT       53870 a 34436 c 33100 g 47270 t 1303 others
ORIGIN

Query Match      69.7%; Score 20.2; DB 2; Length 169979;
Best Local Similarity 88.0%; Pred. No. 92;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 cctgcagccatgctagcccccgc 29
      |||||
Db 73881 TCTGCACCCATGCTGCGCAGCCGC 73905

RESULT 5
LOCUS      AC010811      165412 bp      DNA      linear      HTG 09-FEB-2002
DEFINITION Homo sapiens chromosome 11 clone RP11-3E17 map 11, *** SEQUENCING
IN PROGRESS *** 5 unordered pieces.
ACCESSION   AC010811
VERSION     AC010811.10 GI:18642864
KEYWORDS    HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 165412)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
             Baldwin,J., Barna,N., Beckert,L., Boguski,M., Bouckgeater,B.,
             Brown,A., Castle,A., Collins,S., Collins,S., Collins,M.,
             Cooke,P., Dearfelli,K., Dewar,K., Domino,M., Donnellan,L., Doyle,M.,
             Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
             Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
             Howland,J.C., Johnson,R., Jones,C., Kann,L., Karst,A., Klein,J.,
             Lebecky,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,
             McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
             Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
             Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
             Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
             Tesfaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
             Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
             Direct Submission
             Submitted (23-SEP-1999) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
             On Feb 9, 2002 this sequence version replaced gi:18543112.
             All repeats were identified using RepeatMasker:
             Smit, A.F.A. & Green, P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             ----- Genome Center
             Center: Whitehead Institute/ MIT Center for Genome Research
             Center code: WITB
             Web site: http://www-seq.wi.mit.edu
             ----- Project Information
             Contact: sequence_submissions@genome.wi.mit.edu
             Center project name: L2720
             Center clone name: 3_E_17
             -----
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 5 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will

```

```

* be preserved.
* 1 12794: contig of 12794 bp in length
* 12795 12894: gap of 100 bp
* 12895 68988: contig of 56094 bp in length
* 68989 69088: gap of 100 bp
* 69089 133993: contig of 64903 bp in length
* 133992 134091: gap of 100 bp
* 134092 162167: contig of 28076 bp in length
* 162168 162267: gap of 100 bp
* 162268 165412: contig of 3145 bp in length.
             Location/Qualifiers
             source
               1..165412
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="11"
               /map="11"
               /clone="RP11-3E17"
               /clone_1lb="RP11-3E17"
BASE COUNT    48029 a 30561 c 34765 g 51633 t 424 others
ORIGIN

Query Match      69.0%; Score 20; DB 2; Length 165412;
Best Local Similarity 82.1%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 atatctgcagccatgctagcccccgc 29
      |||||
Db 29415 AGATCTGCACCCATGCTGCGCCCTGC 29442

RESULT 6
LOCUS      AC068549      184018 bp      DNA      linear      HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-559110, WORKING DRAFT
SEQUENCE     AC068549.2 GI:7940407
KEYWORDS    HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 184018)
AUTHORS    Waterston,R.H.
             The sequence of Homo sapiens clone
             unpublished
             2 (bases 1 to 184018)
REFERENCE   Waterston,R.H.
             Direct Submission
             Submitted (03-MAY-2000) Genome Sequencing Center, Washington
             University School of Medicine, 4444 Forest Park Parkway, St. Louis,
             MO 63108, USA
             On May 19, 2000 this sequence version replaced gi:7684582.
             ----- Genome Center
             Center: Washington University Genome Sequencing Center
             Center code: WUGSC
             Web site: http://genome.wustl.edu/gsc/index.shtml
             ----- Project Information
             Contact: sequence_submissions@genome.wi.mit.edu
             Center project name: H_NH0559110
             ----- Summary Statistics
             -----
             Sequencing vector: M13; 100%
             Sequencing vector: plasmid; 0%
             Chemistry: Dye-Primer ET; 100% of reads
             Chemistry: Dye-terminator Big Dye; 0% of reads
             Assembly program: Phrap; version 0.990319
             Consensus quality: 175110 bases at least Q40
             Consensus quality: 178766 bases at least Q20
             Insert size: 204000; agarose-fp
             Insert size: 182018; sum-of-contigs
             Quality coverage: 4.34 in Q20 bases; agarose-fp

```


Quality coverage: 4.93 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1630: contig of 1630 bp in length
* 1631 1730: gap of unknown length
* 1731 3332: contig of 1602 bp in length
* 3333 3432: gap of unknown length
* 3433 5871: contig of 2439 bp in length
* 5872 5971: gap of unknown length
* 5972 8777: contig of 2805 bp in length
* 8777 8877: gap of unknown length
* 8877 11579: contig of 2702 bp in length
* 11579 15035: contig of 3357 bp in length
* 15036 15135: gap of unknown length
* 15136 17606: contig of 2471 bp in length
* 17607 17707: gap of unknown length
* 17707 22355: contig of 4649 bp in length
* 22356 22455: gap of unknown length
* 22456 26204: contig of 3749 bp in length
* 26205 31274: gap of unknown length
* 31275 31374: contig of 4970 bp in length
* 31375 39128: gap of unknown length
* 39129 39228: contig of 7754 bp in length
* 39229 46971: contig of 7743 bp in length
* 46972 47071: gap of unknown length
* 47072 57776: contig of 10705 bp in length
* 57777 57876: gap of unknown length
* 57877 66843: contig of 8967 bp in length
* 66844 66943: gap of unknown length
* 66944 78240: contig of 11297 bp in length
* 78241 78340: gap of unknown length
* 78341 93291: contig of 14951 bp in length
* 93292 93391: gap of unknown length
* 93392 110044: contig of 16653 bp in length
* 110045 110144: gap of unknown length
* 110145 124100: contig of 13956 bp in length
* 124101 124200: gap of unknown length
* 124201 140947: contig of 16747 bp in length
* 140948 141047: gap of unknown length
* 141048 158424: contig of 17377 bp in length
* 158425 184018: gap of unknown length
* 184018 25494 bp in length.
FEATURES
source
1. 184018
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-559110"
1. 1630
misc_feature
1731..3332
/note="assembly_name:Contig9"
misc_feature
clone_end:Sp6
/note="assembly_name:Contig10"
vector_side:left"
3433..5871
/note="assembly_name:Contig11"
5972..8776
/note="assembly_name:Contig12"
8877..11578
/note="assembly_name:Contig13"
11579..15035
/note="assembly_name:Contig14"
15136..17606
/note="assembly_name:Contig15"
```

```
misc_feature 17707..22355
/note="assembly_name:Contig16"
misc_feature 22456..26204
/note="assembly_name:Contig17"
misc_feature 26305..31274
/note="assembly_name:Contig18"
misc_feature 31375..39128
/note="assembly_name:Contig19"
misc_feature 39229..46971
/note="assembly_name:Contig20"
misc_feature 47072..57776
/note="assembly_name:Contig21"
misc_feature 57877..66843
/note="assembly_name:Contig22"
misc_feature 66944..78240
/note="assembly_name:Contig23"
misc_feature 78341..93291
/note="assembly_name:Contig24"
misc_feature 93392..110044
/note="assembly_name:Contig25"
misc_feature 110145..124100
/note="assembly_name:Contig26"
misc_feature 124201..140947
/note="assembly_name:Contig27"
misc_feature 141048..158424
/note="assembly_name:Contig28"
misc_feature 158525..184018
/note="assembly_name:Contig29"
BASE COUNT 54949 a 36654 c 37232 g 53175 t 2008 others
ORIGIN
```

Query Match 69.0% Score 20: DB 2: Length 184018;

Best Local Similarity 82.1% Pred: No. 1.1e+02;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 atattcgcagcatgctagcccccgc 29

Db 63891 AGATCTCAGCAGCATGACTTGGCCCTGC 63918

```
RESULT 7
AC087279 185070 bp DNA linear HTG 13-OCT-2001
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-559110 map 11.*** SEQUENCING
IN PROGRESS ***, 4 unordered pieces.
ACCESSION
AC087279 GI:16117572
VERSION
AC087279.6
KEYWORDS
HTG; HTGS_PHASE1; HTGS_FULUTOP; HTGS_ACTIVEPIN.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 185070)
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Bouchehater,B., Brown,A.,
Cammarata,J., Campopiano,A., Choquet,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glade,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Huime,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., Lahocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meidrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunhahng,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
```


260689 260871: gap of unknown length
 * 260872 260884: contig of 7213 bp in length
 * 268085 268267: gap of unknown length
 * 268268 275138: contig of 6871 bp in length
 * 275139 275321: gap of unknown length
 * 275322 281763: contig of 6442 bp in length
 * 281764 281946: gap of unknown length
 * 281947 288236: contig of 6230 bp in length
 * 288237 288419: gap of unknown length
 * 288420 294677: contig of 6258 bp in length
 * 294678 294860: gap of unknown length
 * 294861 301014: contig of 6154 bp in length
 * 301015 301197: gap of unknown length
 * 301198 306779: contig of 5582 bp in length
 * 306780 306962: gap of unknown length
 * 312501 312501: contig of 5539 bp in length
 * 312502 312684: gap of unknown length
 * 317929 317929: contig of 5245 bp in length
 * 317930 318112: gap of unknown length
 * 318113 323028: contig of 4916 bp in length
 * 323029 323211: gap of unknown length
 * 323212 328092: contig of 4881 bp in length
 * 328093 328275: gap of unknown length
 * 328276 332714: contig of 4439 bp in length
 * 332715 332897: gap of unknown length
 * 332898 336810: contig of 3913 bp in length
 * 336811 336993: gap of unknown length
 * 336994 340782: contig of 3789 bp in length
 * 340783 340965: gap of unknown length
 * 340966 344707: contig of 3742 bp in length
 * 344708 344890: gap of unknown length
 * 344891 348401: contig of 3511 bp in length
 * 348402 348584: gap of unknown length
 * 348585 352048: contig of 3464 bp in length
 * 352049 352231: gap of unknown length
 * 352232 352375: contig of 144 bp in length.

FEATURES
 source
 1. 352375
 Location/Qualifiers

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-93N5"

BASE COUNT 92469 a 84017 c 82996 g 86234 t 6659 others
 ORIGIN

Query Match 68.3% Score 19.8; DB 2; Length 110000;
 Best Local Similarity 91.3% Pred. No. 1.4e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 tatctgcagcagctagctagcc 25
 DB 60510 TATCTGCAGCCTAGCTAGCCCC 60488

RESULT 9
 AC094266/c 115062 bp DNA linear HTG 20-DEC-2001
 LOCUS Rattus norvegicus clone CH230-3L5, *** SEQUENCING IN PROGRESS ***
 DEFINITION 47 unordered pieces.
 AC094266
 VERSION AC094266.2 GI:17940982
 KEYWORDS HTG; HTGS-PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE
 AUTHORS
 1 (bases 1 to 115062)
 Murny, D.M., Adams, C., Adlo-Oduola, B., Ali-osman, F.R., Allen, C.,
 Albrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbara, J.,
 Benton, J., Bilmage, K., Blankenburg, K., Bonnin, D., Bouck, J.,
 Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
 Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
 Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
 Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
 Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Diaper, H.,
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flaeg, N., Ford, J.,
 Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
 Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,
 Hamilton, K., Harris, C., Harris, K., Hart, M., Haykal, P., Hayes, A.,
 Hernandez, J., Hernandez, O., Hodgson, A., Hughes, M., Holloway, C.,
 Hollins, B., Homsl, F., Howard, S., Huber, J., Huylk, S., Hume, J.,
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivel, S.,
 Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovach, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
 Louised, H., Lozano, R.J., Lu, X., Lueker, A., Lueker, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
 Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwuo, S.,
 Oguri, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, U., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubo, I., Rolfe, M.,
 Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshbari, N.,
 Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Swalek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinsón, R.,
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Wallington, S., Williams, G., Williamson, A., Wlecyk, R., Wooden, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,
 Welsch, G. and Gibbs, R.

Direct Submission
 Unpublished
 2 (bases 1 to 115062)
 Worley, K.C.

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced g1:15636609.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GAHL

Center clone name: CH230-3L5

Summary Statistics

Assembly program: Phrap; version 0.9903299First call to

findhaplolist

Consensus quality: 96225 bases at least Q40

Consensus quality: 105112 bases at least Q30

Consensus quality: 112446 bases at least Q20

Estimated insert size: 93083; sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agatose-fp estimation

Quality coverage: 1.2x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 47 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 12330: contig of 12330 bp in length

12331 12430: gap of unknown length

```

* 12431 18306: contig of 5876 bp in length
* 18307 18406: gap of unknown length
* 18407 22841: contig of 4435 bp in length
* 22842 22942: gap of unknown length
* 22942 26470: contig of 3529 bp in length
* 26471 32225: gap of unknown length
* 32225 32325: contig of 5656 bp in length
* 32325 33350: gap of unknown length
* 33350 35451: contig of 3024 bp in length
* 35451 38659: contig of 3209 bp in length
* 38659 42445: gap of unknown length
* 42445 42545: contig of 3687 bp in length
* 42545 45127: gap of unknown length
* 45127 45227: contig of 2581 bp in length
* 45227 46887: gap of unknown length
* 46887 46987: contig of 1660 bp in length
* 46987 50671: gap of unknown length
* 50671 50771: contig of 3684 bp in length
* 50771 52400: gap of unknown length
* 52400 52500: contig of 1629 bp in length
* 52500 55326: gap of unknown length
* 55326 55426: contig of 2826 bp in length
* 55426 58414: gap of unknown length
* 58414 58514: contig of 2988 bp in length
* 58514 61460: gap of unknown length
* 61460 61560: contig of 2946 bp in length
* 61560 63526: gap of unknown length
* 63526 63626: contig of 1966 bp in length
* 63626 65627: gap of unknown length
* 65627 65784: contig of 2158 bp in length
* 65784 65884: gap of unknown length
* 65884 68062: gap of unknown length
* 68062 68162: contig of 2178 bp in length
* 68162 69962: gap of unknown length
* 69962 70062: contig of 1800 bp in length
* 70062 72735: gap of unknown length
* 72735 72835: contig of 2674 bp in length
* 72835 74715: gap of unknown length
* 74715 74816: contig of 1880 bp in length
* 74816 76796: gap of unknown length
* 76796 76896: contig of 1980 bp in length
* 76896 78077: gap of unknown length
* 78077 80644: gap of unknown length
* 80644 80744: contig of 2468 bp in length
* 80744 82356: gap of unknown length
* 82356 82456: contig of 1612 bp in length
* 82456 84035: gap of unknown length
* 84035 84135: contig of 1579 bp in length
* 84135 86753: gap of unknown length
* 86753 86853: contig of 2618 bp in length
* 86853 88307: gap of unknown length
* 88307 88407: contig of 1454 bp in length
* 88407 89643: gap of unknown length
* 89643 89743: contig of 1236 bp in length
* 89743 91020: gap of unknown length
* 91020 91120: contig of 1277 bp in length
* 91120 92796: gap of unknown length
* 92796 94248: contig of 1576 bp in length
* 94248 94348: gap of unknown length
* 94348 95368: contig of 1452 bp in length
* 95368 95468: gap of unknown length
* 95468 96603: contig of 1020 bp in length
* 96603 96703: gap of unknown length
* 96703 98728: contig of 2025 bp in length
* 98728 98828: gap of unknown length
* 98828 99950: gap of unknown length
* 99950 100051: contig of 1122 bp in length
* 100051 101229: gap of unknown length
* 101229 101329: contig of 1179 bp in length
* 101329 102503: gap of unknown length
* 102503 102503: contig of 1174 bp in length

```

FEATURES

```

* 102504 102603: gap of unknown length
* 102604 103796: contig of 1193 bp in length
* 103797 103896: gap of unknown length
* 103897 105091: contig of 1195 bp in length
* 105092 105191: gap of unknown length
* 105192 106200: contig of 1009 bp in length
* 106201 106301: gap of unknown length
* 106301 107720: contig of 1420 bp in length
* 107721 107820: gap of unknown length
* 107821 109208: contig of 1388 bp in length
* 109209 109308: gap of unknown length
* 109309 110945: contig of 1637 bp in length
* 110946 111045: gap of unknown length
* 111046 112361: contig of 1316 bp in length
* 112362 112461: gap of unknown length
* 112462 113735: contig of 1274 bp in length
* 113736 113835: gap of unknown length
* 113836 115062: contig of 1227 bp in length.

```

SOURCE

```

1. 115062
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-315"

```

```

BASE COUNT 33325 a 20155 c 21211 g 35735 t 4636 others
ORIGIN

```

Query Match

```

Best Local Similarity 91.3% Score 19.8; DB 2; Length 115062;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY

```

4 atctgcagcagtgtagccccc 26
1111111111111111111111
Db 76679 AACTGCAGCAGCAGCTAGGCCCC 76657

```

RESULT 10

```

AC079643/c 151733 bp DNA linear HTG 04-JUL-2001
LOCUS
DEFINITION

```

ACCESSION

```

AC079643 GI:14595775
VERSION

```

KEYWORDS

```

HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE

```

ORGANISM

```

Mus musculus.
REFERENCE

```

AUTHORS

```

1 (bases 1 to 151733)
McCombie,W.R., Baker,V.P., Bahret,A., Yang,C., Balija,V.,

```

TITLE

```

Decilia,N.N., de la Bastide,M., Kuit,K., King,L., Kirchoff,R.A.,

```

JOURNAL

```

Miller,B., Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R.,

```

COMMENT

```

Rodriguez,S., Santos,L., Shah,R.S., Spiegel,L.A., Palmer,L.,

```

JOURNAL

```

Vill.M.D. and Zuttervern,T.
Mouse Genomic Sequence

```

REFERENCE

```

Unpublished
2 (bases 1 to 151733)
McCombie,W.R.

```

AUTHORS

```

Direct Submission
Submitted (06-SEP-2000) Lita Annenberg Hazen Genome Sequencing

```

JOURNAL

```

Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA

```

COMMENT

```

On Jul 4, 2001 this sequence version replaced gi:10280742.
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: http://www.cshl.org/genseq
Contact: mcombie@cschl.org
----- Project Information
Center project name: RP23-79H19
Center clone name: RP23-79H19
-----

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

24049: contig of 24049 bp in length
24050: gap of unknown length
24489: contig of 16895 bp in length
41384: gap of unknown length
41823: contig of 12958 bp in length
54780: gap of unknown length
55220: contig of 11991 bp in length
67210: contig of unknown length
67649: gap of unknown length
67650: contig of 10300 bp in length
77950: gap of unknown length
78388: contig of 10230 bp in length
88618: gap of unknown length
89053: contig of 7638 bp in length
96693: gap of unknown length
96994: gap of unknown length
97132: contig of 6436 bp in length
103567: contig of 6329 bp in length
104005: gap of unknown length
104006: contig of 6329 bp in length
110335: gap of unknown length
110773: contig of 6060 bp in length
116833: gap of unknown length
117271: contig of 5745 bp in length
123015: gap of unknown length
123016: gap of unknown length
123453: contig of 5590 bp in length
129043: gap of unknown length
129481: gap of unknown length
129482: contig of 4571 bp in length
134053: gap of unknown length
134490: gap of unknown length
138898: contig of 4408 bp in length
139336: gap of unknown length
139337: contig of 4122 bp in length
143458: gap of unknown length
143459: gap of unknown length
147887: contig of 3991 bp in length
147888: gap of unknown length
148325: gap of unknown length
151733: contig of 3408 bp in length.

FEATURES

SOURCE

1. 151733
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-79H19"
/clone_id="RPC1-23"

BASE COUNT 33878 a 37878 c 38717 g 34033 t 7227 others
ORIGIN

Query Match 68.3% Score 19.8; DB 2; Length 151733;
Best Local Similarity 91.3% Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 tatctgcagcatggttagccccc 25

Db 94130 TATCTGCAGCATGCTAGCGCC 94108

RESULT 11
AF321096/c 1560 bp DNA linear PLN 15-MAR-2001
LOCUS
DEFINITION Pichia pastoris PR-aminimidazoleuccinocarbamide synthase (ADE1)
ACCESSION AF321096
VERSION AF321096.1 GI:12744763
KEYWORDS
SOURCE Pichia pastoris.
ORGANISM Pichia pastoris
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.

REFERENCE 1 (bases 1 to 1560)

AUTHORS

Lln Cereghino,G.P., Lln Cereghino,J., Sunga,A.J., Johnson,M.A.,
Lim,M., Gleeson,M.A.G. and Cregg,J.M.

TITLE

New selectable marker/auxotrophic host strain combinations for
molecular genetic manipulation of Pichia pastoris

JOURNAL

Gene 263 (1-2), 159-169 (2001)

PUBMED

11223254

REFERENCE

2 (bases 1 to 1560)
Lln Cereghino,G.P., Lln Cereghino,J., Sunga,A.J., Johnson,M.A.,
Lim,M., Gleeson,M. and Cregg,J.M.

TITLE

Direct Submission
Submitted (14-NOV-2000) Biological Sciences, University of the
Pacific, Stockton, CA, USA

FEATURES

Location/Qualifiers
1. 1560

SOURCE

/organism="Pichia pastoris"
/db_xref="taxon:4922"

CDS

gene
/product="PR-aminimidazoleuccinocarbamide synthase"
/gene="ADE1"
371..1285
/gene="ADE1"
/function="biosynthetic enzyme necessary for production of
adenine"
/note="adel mutant has pink phenotype"
/codon_start=1
/product="PR-aminimidazoleuccinocarbamide synthase"
/protein_id="AAK06766.1"
/db_xref="GI:12744764"

BASE COUNT

494 a 285 c 329 g 452 t

ORIGIN

Query Match 67.6% Score 19.6; DB 8; Length 1560;
Best Local Similarity 84.6% Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 atactgcagcatggttagccccc 27

Db 1218 ATTTCGACATCCATGCGCTACGCCCTC 1193

RESULT 12

AX175474 11358 bp DNA linear PAT 03-JUL-2001

LOCUS

Sequence 3 from Patent WO0144443.

DEFINITION

AX175474

ACCESSION

AX175474.1 GI:14598816

VERSION

AX175474.1

KEYWORDS

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 11358)
White,J.A., Petkovich,P.M., Jones,G. and Ramshaw,H.

AUTHORS

Cytochrome p450ra1-2 and related proteins
Patent: WO 0144443-A 3 21-JUN-2001;

JOURNAL

Cytochrome Inc. (CA)

FEATURES

Location/Qualifiers
1. 11358

SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT

2654 a 2894 c 3040 g 2770 t

ORIGIN

gene
 CDS
 complement(6490..7440)
 /gene="MT2903"
 complement(6490..7440)
 /note="Identified by match to protein family HMM"
 /codon_start=1
 /transl_table=11
 /product="DHH family protein"
 /protein_id="AAK47229.1"
 /db_xref="GI:13882677"
 /translation="MDAVGAALILSAAARGVYCHYRPADTIATGALALVLDGCGK
 RVEVSRAPATLPESLRSPECHILVPEVRKRVDLVVYDIPSYRGLAGDUTDS
 GRELVLIDHNASNDLFTGTANFIDPSASTTTTMAEILDAMGKLPIDPRVAHITVGLAN
 DTGFRMAVSQVGRILARLVEIGVNDATVSTRTLSDSHPEFMPLPLSIVLSAQCLVSR
 TVNLAAYASGVGGGSHRLAAGYTTGTSIDDAVAVSLRAALG"
 complement(7475..8026)
 /gene="MT2904"
 complement(7475..8026)
 /gene="MT2904"
 /note="similar to SP:P32731 GB:Z18631 PID:580902
 GB:U009126; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="ribosome-binding factor A"
 /protein_id="AAK47230.1"
 /db_xref="GI:13882678"
 /translation="MADAAARARRLAKRIIAIVASAIEYEIKPGLAGVTTITDKATAD
 LHDIVYTVYWGRTLHDEPNCAGAAALERAKGYLRKVGAGTVRFPLTFITDITIT
 SDSVHRDELLEARARADADLARVYGAKPAGEADPYRDNQSVASQAPAGSLGRTSD
 GPEAVEAPLTCCGPTGDDDRPKE"
 complement(8026..10728)
 /gene="MT2905"
 complement(8026..10728)
 /gene="MT2905"
 complement(8026..10728)
 /note="similar to SP:P04766 GB:X04399 PID:39954;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="translation initiation factor IF-2"
 /protein_id="AAK47231.1"
 /db_xref="GI:13882679"
 /translation="MAAGKARVHELAKELGTSEVILARLSEGEFVKSASTVEADY
 ARRLRSRSGSKSPAPAKGTAKSPCKGDKSLDKRLDAIDMAANGKATAPAKAAS
 GGAIVPTTPPAEPPTAVPPSQAPHPCMAPRGPPVPKPCGIRTPRGNNPFSSA
 QPADRPIDPPAPRPGTARPEVPRGASPCSMPRPGGAVGAGRPARGARPGGRGCG
 APGAGRSDAGGPNRGGVGAAPTGTGRGRGGGGGPGCGGGAAGAFGRPGAPRRR
 GRKSRKROEYDSMQAPVYGVVRLPHNGNGLTILAGASLSIDPADRIDAPALVQDA
 LFNENGEVTTAQSGDELTLELISEMNYNOVAVPEDEDRLESPDISYGEDGGEE
 DLYQRPVYVYMGVHDGKTRILDTIKARNAEAGGITQHGATQVADYDQSOBL
 ITFDTPGHEAFTEMRARCAKNDIILVYAADDGCVAPQTEALINHQADVPIVYVA
 NKIDKESADRAKIRNGLTLEYGLVPEEFGDMFDISAKGTINLEAEVALLPADAA
 LDLEANPDMEQVAIEFAHLDRGRGPAVYLVQGTLRVGDSVVAAGVAGVRVMDVDE
 HGEDEVALPSRPVQIGTSPGAGNGLFVDEDRLARQIADRSARKNALAAERSR
 KLSLEDLDSALKETSOQLNLKGDNAQVTEALEALMGIDYDEVVLRYVIDRGGVIG
 TETVNLASASDAVITGFNVRAEKAETELSSRECEVLRYSYVQALDEIQALRGILL
 KPIYENOLDGAEIRALFRSRKVGTLICVTGSGVMRNMARLLRDNIVYAENLSIA
 SLREKNDVTEVRGPECGCTIGYADIKREGDVIESTELYOKERA"
 complement(10814..11113)
 /gene="MT2906"
 /gene="MT2906"

RESULT	14
LOCUS	MSGY414A/c
DEFINITION	40121 bp DNA linear BCT 03-DEC-1996
ACCESSION	AB000007
VERSION	AD000007.1 GI:1702972
KEYWORDS	
SOURCE	
ORGANISM	Mycobacterium tuberculosis (clone: y414a) ds-DNA.
REFERENCE	Mycobacterium tuberculosis Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex. 1 (bases 1 to 40121) Du,L.
AUTHORS	Direct Submission
TITLE	Submitted (11-OCT-1996) L. Du, Genome Therapeutics Corporation, 100
JOURNAL	Beaver Street, Waltham, MA, USA, 02154 du@etric.com
COMMENT	GSDB:s:1004719
FEATURES	Location/Qualifiers
source	1..40121 /organism="Mycobacterium tuberculosis" /db_xref="taxon:1773" /clone="y414a"
BASE COUNT	7156 a 12828 c 13251 g 6885 t 1 others
ORIGIN	

Query Match	67.6%	Score 19.6	DB 1	Length 40121
Best Local Similarity	84.6%	Pred. No. 1.7e+02		
Matches 22:	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

OY	4	atctgacgcatggctagcccccgc 29
Db	5823	ATCCGACACCACGGCGAAGCCGCCGC 5798

RESULT 15	LOCUS	MYC116B7	43330 bp	DNA	linear	BCT 03- AUG-2001
DEFINITION	MYCOBACTERIUM TUBERCULOSIS H37RV	complete genome; segment 123/162.				
ACCESSION	Z81331	AL123456				
VERSION	281331.1	GI:3261650				
KEYWORDS						
SOURCE	MYCOBACTERIUM TUBERCULOSIS H37RV.					
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriineae; Mycobacteriineae; Mycobacterium tuberculosis complex.					
REFERENCE	1 (bases 1 to 43430)					
AUTHORS	Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltham, D., Gentles, S., Hamlin, N., Holtley, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Randle, J., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Squares, R., Sulston, J.E., Taylor, K., Whitehead, S. and Barrett, B.G.					
TITLE	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence					
JOURNAL	Nature. 393 (6685), 537-544 (1998)					
MEDLINE	96293587					
REFERENCE	2 (bases 1 to 43430)					
AUTHORS	Parkhill, J.					
TITLE	Direct Submission					
JOURNAL	Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk					
COMMENT	On Jun 27, 1998 this sequence version replaced gi:11648878.					

Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.

(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submitters but the old gene designations are in brackets after the new gene numbers. Gene prediction was based on a Hidden Markov Model of 17 genes implemented in TParse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C. CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, ggg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

```
FEATURES
source      1..43430
             /organism="Mycobacterium tuberculosis H37Rv"
             /strain="H37Rv"
             /db_xref="taxon:83332"
             <1..507
             /note="fragment designated v002. Does not represent a
             physical clone"
             58..113
             /note="hairpin loop with 38 bp stem"
             /complement(133..696)
             /gene="lppv"
             /complement(133..696)
             /gene="lppv"
             /note="Rv2796c, (MTY002.61c-MTCY16B7.47), len 187;
             Probable lipoprotein similar to MTCY159.13c (220 aa) fasta
             score:opt: 169 z-score: 206.2 E(): 0.00058; 22.4%
             identity in196 aa overlap"
             /codon_start=1
             /transl_table=11
             /product="lppv"
             /protein_id="CAB03673.1"
             /db_xref="GI:3261651"
             /db_xref="SPTREMBL:P71655"
             /transl_table="MRPPTMALLALVCVMATGCGPSGSGHGRAGEEPLSPKVALEEN
             PLAKRPLEDAKQYRAVATOLAITALVPGITWRDMDTNGCGGEYMTAKAY
             FMIVSGSPTRDDKMLQAVQIVKGVDFGATGCGVAKKPADVDYFAGGVEFKCS
             TOKAAVLTAOSDCRISTRTDKPSPPT"
             382..>43430
             /organism="Mycobacterium tuberculosis H37Rv"
             /strain="H37Rv"
             /db_xref="taxon:83332"
             /clone="Y16B7"
             /complement(696..2384)
             /gene="Rv2797c"
             /complement(696..2384)
             /gene="Rv2797c"
             /note="Rv2797c, (MTCY16B7.46), len: 562. Unknown,
             contains 500120 lipases, serine active site, similarity to
             283863|MTCY159.14 Mycobacterium tuberculosis cosmid (403
             aa) opt:537 z-score: 530.8 E(): 4.8e-22; 40.8%
             identity in 292aa overlap"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein Rv2797c"
             /protein_id="CAB03649.1"
             /db_xref="GI:1648924"
             /db_xref="SPTREMBL:P71654"
             /transl_table="MPLTVADIDRMNNAQAVREVFHAASAREVTFEASROLALSTFA
             NSGGTAAEAHHNAGIRDLDAHGNDAVAAADRADGIVKVOSELALHAAHA
             AETLIDALINRVYPIGIRSTENQARTAKTELDIAEDLAIKANAENDEEDQLSV
             MADGAPIPADSGPVPGELPTQULASDNERELKEERARLQAHLERLQAEYDQLSV
             RAARDYHNGILDDADAGRLAALDELISAANGRLGELDAVDALSRAPETYLQIQPE
             DPNOQVLAAVAGNDPTAANVSIVPGVSTGRGALPCWTEARDLRESEYIRLOQLAE
             KPASVNTIAMGYPHPNPIDGVSAGDLMQMTWDGAAHAGADLSRYLQOVRRANNNG
             HLTGLSHGSGSLASLALODLDQASNAHPNDVYFYFSGELTSPNQLGIDRHQATYMS
             QAFHDLITNLVAPLAPLHGWGLDPTLTPGFTELSSQAGFDGPGIMRWGYAHGDPYS
```

```
misc_feature
             /complement(1176..1205)
             /gene="Rv2797c"
             /note="PS00120 Lipases, serine active site"
             /complement(2388..2714)
             /gene="Rv2798c"
             /complement(2388..2714)
             /gene="Rv2798c"
             /note="Rv2798c, (MTCY16B7.45), len: 108. Some similarity
             to Rv0965c(MTCY10D7.09) (139 aa, E(): 1.6e-06, 38.9%
             identity in 90 aa overlap)"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein Rv2798c"
             /protein_id="CAB03648.1"
             /db_xref="GI:1648923"
             /db_xref="SPTREMBL:P71653"
             /transl_table="MFOISPEOMHSAQVTTQGEGLAVHGLSSDYRMQAQGRNOGA
             SAMALNAKMDMDLSDASRALTRIGDHFGLQEAIAIOHAAEAERQALQGVSADEV
             AGPRGV"
             2845..3474
             /gene="Rv2799"
             2845..3474
             /gene="Rv2799"
             /note="Rv2799, (MTCY16B7.44c), len: 209. Unknown"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein Rv2799"
             /protein_id="CAB03647.1"
             /db_xref="GI:1648922"
             /db_xref="SPTREMBL:P71652"
             /transl_table="MTTPGKPPRAGGVFTFRRLIGLGLALTAAVVYGVGMGIP
             PAPTGDDAVQLRSTAPMSYTMKSPVATTPDPSPPCRDIPDVIQRLGLATPPEA
             EBLGRCHFADGNQMAVEPIIMRTYQTLPPDAIETTLIGHRAAQVWRKPTVHNSFM
             YSSCWTFKTSYGVCIQSLFYSVSPVDVDCSTNLQRANDLVYYRF"
             3483..5142
             /gene="Rv2800"
             3493..5142
             /gene="Rv2800"
             /note="Rv2800, (MTCY16B7.43c), len: 549. Some similarity
             to G216374 GLUTARYL 7-ACA ACYLASE precursor (634 aa)
             fasta scores, opt: 362. E(): 1e-14, (26.7% identity in 589
             aa overlap) and to D909031D90903.22 Synchocystis sp:
             PCC6803 complete (535 aa) opt: 455 z-score: 494.2 E():
             5.3e-20; 29.8% identity in 503 aa overlap. Some similarity
             to Rv1835c, (MTCY1A11.08)"
             /codon_start=1
             /transl_table=11
             /product="hypothetical protein Rv2800"
             /protein_id="CAB03646.1"
             /db_xref="GI:1648921"
             /db_xref="SPTREMBL:P71651"
             /transl_table="MSTTSARPERPKRALTGVNGGOALGGLDLPATTVYGVHR
             VPMRDVQVLAHYADHAPATSOVCTLLVGGVGRPSVYFAITVAAAGVHVLSVR
             GTGSGGVPERPNEAADGADIVAMLRDEQWFTGRTGTLPLTGLTQALLDPPPE
             LAAAVTGVPHDRASVWGTSFTVNDVDFGWSLVSHQEDPGRIAGTRQVLRRAVA
             RTAATLPLGESSARTLLGTGAPNFSVNEHPTDRDPMDFRPAALDRQVLPVLVGG
             MODIFLQTLQOQYRHLRDRGVHVALTVGPTHLQMLTKGLATARSGLDMLDAHLGA
             PALRSPVRYVTEVGGCWRHLPDMPAPTAETBAVYLOGGRGLGASAPASGTPAFERHAP
             ADPPTTGGPLSSNGGYRDSRLATRADVCTGAPLITDLGVHNPVETVHSSDN
             PYDVYVRVSEVDYAKGRSVSGYRRLGDAPLVYELDAIHRRAADSRIVYLLAG
             SWPRTYARNLGTPEPLTIGROLKPATHAVHFGSRLLPLVG"
             /complement(5244..5600)
             /gene="Rv2801c"
             /complement(5244..5600)
             /gene="Rv2801c"
             /note="Rv2801c, (MTCY16B7.42), len: 118. Similar to B.
             subtilis sp|P96622|P96622 YDCE PROTEIN (116 aa) fasta
             scores, opt: 194 z-score: 275.5 E(): 8e-08; 33.3%
             identity in 117 aa overlap. PEMK ECOLI P13976 pemk protein
             (133 aa), fasta scores, opt: 153. E(): 0.00055 (35.6%
             identity in 118 aa overlap), also MTCY39.28 (36.8%
             identity in 117 aa overlap)"
```


gene
CDS
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv2801c"
/protein_id="CAB03645.1"
/db_xref="GI:1648920"
/db_xref="SPTREMBL:P71650"
/translation="MMRGEITWQYDLPARGSEANNQRPAYVVSNDRAATATRLGRG
VITVVPVTSNIATKYPFQVLLSATTGQLQVDCKAQAEQIRSIATERLLRP1GVSAAE
LAQLDEALKHLIDLMS"
complement(5857..6900)
/gene="Rv2802c"
complement(5857..6900)
/note="Rv2802c, (MTCY16B7.41), len: 347. Unknown"
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv2802c"
/protein_id="CAB03674.1"
/db_xref="GI:3261652"
/db_xref="SPTREMBL:P71649"
/translation="MARQPLEQRVARAAQAALARQRFVSAIDVLLGLGWLAPSHVDW
OY 4 atctgcagccatgctagggcccccgcg 29
||| ||||| ||| ||||| |||||
db 38302 ATCCGACGCCACGCGACGCCGCCGCG 38327

Search completed: August 31, 2002, 18:03:50
Job time: 15683 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:43:18 ; Search time 824.08 Seconds
(without alignments)
60.420 Million cell updates/sec

Title: US-09-810-861b-1

Perfect score: 29

Sequence: 1 gatatctgcagccatgctagggccccgc 29

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /N_Geneseq_032802.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	29	24	AA517545 Human acetylcholin
2	19.6	67.6	11358	22	AAH22434 P450RAI-2 containi
3	19.2	66.2	375	22	AAF66350 Novel human polynu
4	19.2	66.2	657	22	AAH52128 Human AFP protein
5	19.2	66.2	1921	23	AA576387 DNA encoding novel
6	19.2	66.2	2390	22	AA545000 cDNA encoding nove
7	19.2	66.2	3023	23	AA577493 DNA encoding novel
8	19.2	66.2	3684	23	AA580007 DNA encoding novel
9	19	65.5	934	22	AA536719 Human musculoskele

ALIGNMENTS

RESULT 1

AA517545
ID AA517545 standard; DNA; 29 BP.

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

XX AC AA517545;

PI Mor T, Soreq H, Arntzen C, Mason H;
 DR WPI; 2002-055120/07.
 XX
 XX Production of a transgenic plant which contains a polynucleotide that
 PT encodes a human acetylcholinesterase which upon purification is
 PT effective against acetylcholinesterase poisoning -
 XX
 XX Claim 11: Page 4; 42pp; English.
 XX
 CC The invention relates to a method of producing a transgenic plant which
 CC contains a polynucleotide that encodes human acetylcholinesterase (ACHE)
 CC which upon purification is effective against acetylcholinesterase
 CC poisoning. The method is used for treating a victim of
 CC acetylcholinesterase poisoning by administering a therapeutic amount of
 CC a physiologically active human acetylcholinesterase expressed in plant
 CC tissue. The extensive use of anticholinesterase pesticides with
 CC concurrent accidental poisoning, the threat of chemical warfare and
 CC environmental concerns demand the development of effective, inexpensive
 CC and stage countermeasures and bioremediation solutions. Prior art methods
 CC for treating ACHE poisoning have used the muscarinic receptor antagonist
 CC atropine and oximes to reactivate the organophosphate(OP)-modified ACHE.
 CC The reversible carbamate, pyridostigmine bromide has also been used as a
 CC prophylactic. However, these conventional treatments have limited
 CC effectiveness and serious short and long-term side effects and may result
 CC in significant performance deficits and even permanent brain damage. This
 CC invention permits the utilisation of cholinesterases to counter-act the
 CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
 CC production of the enzymes is cost effective and the product is stable
 CC so that the injected enzymes have the advantage of having a long
 CC half-life. The transgenic form of the enzymes are also easy to purify.
 CC The present sequence represents PCR primer ACHE-Nco used to amplify
 CC human ACHE as described in the method of the invention.
 XX
 SQ Sequence 29 BP; 5 A; 11 C; 8 G; 5 T; 0 other;

Query Match 100.0%; Score 29; DB 24; Length 29;
 Best Local Similarity 100.0%; Pred. No. 0.0066;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatattctgcagccatggtctagggcccccgc 29
 |||||
 DB 1 gatattctgcagccatggtctagggcccccgc 29

RESULT 2
 AAH22434/C
 ID AAH22434 standard; DNA; 11358 BP.
 XX
 AC AAH22434;
 XX
 DT 22-AUG-2001 (first entry)
 XX
 DE P450RAI-2 containing human genomic DNA (clone NH0493L16).
 XX
 KW Cytochrome P450; P450RAI-2; brain; retinoic acid; cancer; dysplasia;
 KW autolimmune; dermatological; cytostatic; antiinflammatory; antiseborrheic;
 KW antipsoriatic; immunosuppressive; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200144443-A2.
 XX
 PD 21-JUN-2001.
 XX
 PF 15-DEC-2000; 2000WO-CA01493.
 XX
 XX 16-DEC-1999; 99US-0171110.
 PR 27-JAN-2000; 2000US-0178314.
 XX
 XX (CVTO-) CYTOCHROMA INC.
 PA
 XX

PI White JA, Petkovich PM, Jones G, Ramshaw H;
 DR WPI; 2001-390242/41.
 DR P-PSDB; AAB85151.
 XX
 XX Novel P450 protein useful for metabolizing retinoic acid for treating
 PT cancer, dysplasia, an autoimmune or dermatological disease -
 XX
 XX Example 1; Page 155-157; 174pp; English.
 XX
 CC The present invention provides a novel all-trans-RA metabolising
 CC cytochrome P450, P450RAI-2, that is predominantly expressed in the
 CC brain. This novel cytochrome P450 is useful for metabolising retinoic
 CC acid in a cell or organism, for screening drugs for their effect of
 CC protein activity, oxidizing a retinoid, screening an agent for its effect
 CC on protein activity. The P450RAI-2 polypeptide, nucleic acid and host
 CC cells containing them are useful for treating cancer, dysplasia, an
 CC autoimmune or dermatological disease. A drug which has an effect on the
 CC expression of P450RAI-2 is used to inhibit retinoic acid metabolism in
 CC the treatment of cancer, actinic keratosis, oral leukoplakia, a secondary
 CC head and/or neck tumour, a non-small cell lung carcinoma, a basal cell
 CC carcinoma, skin cancer, and a premalignancy associated actinic keratosis,
 CC acne, psoriasis, ichthyosis, and/or preferably acute promyelocytic from
 CC leukemia. The present sequence represents a human genomic sequence from
 CC GenBank (Accession#, AC007002 clone name NH0493L16). This sequence
 CC comprises the DNA encoding the novel cytochrome P450, P450RAI-2.
 XX
 SQ Sequence 11358 BP; 2654 A; 2894 C; 3040 G; 2770 T; 0 other;

Query Match 67.6%; Score 19.6; DB 22; Length 11358;
 Best Local Similarity 84.6%; Pred. No. 87;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 atctgcagccatggtctagggcccccgc 29
 |||||
 DB 760 ATCTGCACCCAGGCCAGGCCCCAGC 735

RESULT 3
 AAF66350
 ID AAF66350 standard; cDNA; 375 BP.
 XX
 AC AAF66350;
 XX
 DT 09-APR-2001 (first entry)
 XX
 DE Novel human polynucleotide, SEQ ID NO: 2106.
 XX
 KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 KW breast cancer; lung cancer; cancer detection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200102568-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US18374.
 XX
 PR 02-JUL-1999; 99US-0142310.
 PR 02-JUL-1999; 99US-0142311.
 XX
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;
 XX
 XX WPI; 2001-091805/10.
 DR
 XX

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 850; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
XX dysplasia and hyperplasia.
XX
SQ Sequence 375 BP; 71 A; 140 C; 105 G; 58 T; 1 other;

Query Match 66.2%; Score 19.2; DB 22; Length 375;
Best Local Similarity 87.5%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 atctgcagccatggcctagggccccc 27
|||||
Db 241 atctgcagccatggcctagggccccc 264

RESULT 4
AAH52128
ID AAH52128 standard; cDNA; 657 BP.
XX
AC AAH52128;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human AFP protein encoding cDNA sequence SEQ ID NO:71.
XX
KW Human; secreted protein; secretion; bacterial cell; fungal cell;
KW eukaryotic cell; fusion protein; maltose binding protein;
KW immunoglobulin constant region; polyhistidine tag; ss.
XX
OS Homo sapiens.
XX
PN WO200129221-A2.
XX
PD 26-APR-2001.
XX
PF 20-OCT-2000; 2000WO-US29052.
XX
PR 20-OCT-1999; 99US-0160712.
XX
PS (ZYMO) ZYMOGENETICS INC.
XX
PA Conklin DC, Yee DP;
XX
PI WPI; 2001-300340/31.
XX
DR P-PSDB; AAG81277.
XX
CC Isolated polypeptide for directing secretion of proteins of interest
CC from a host cell including, e.g. bacteria, includes contiguous amino
CC acid residues of polypeptide with specified amino acids -
XX
PS Claim 9; Page 174-176; 617pp; English.
XX
CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
CC to AAG81453. The secreted proteins can be used for directing the

CC secretion of proteins of interest from a host cell including bacteria,
CC fungal cells, and cultured higher eukaryotic cells. The present invention
CC also describes fusion proteins, where a secreted protein of the invention
CC is operably linked via a peptide bond or peptide linker to a second
CC protein selected from the group consisting of maltose binding protein,
CC an immunoglobulin constant region, a polyhistidine tag and a peptide
CC given in AAG81453.
XX
SQ Sequence 657 BP; 132 A; 214 C; 207 G; 104 T; 0 other;

Query Match 66.2%; Score 19.2; DB 22; Length 657;
Best Local Similarity 87.5%; Pred. No. 99;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 atctgcagccatggcctagggccccc 27
|||||
Db 571 atctgcagccatggcctagggccccc 594

RESULT 5
AAS76387
ID AAS76387 standard; cDNA; 1921 BP.
XX
AC AAS76387;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12191.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG12200.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 12191; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1921 BP; 445 A; 558 C; 540 G; 378 T; 0 other;

Query Match 66.2%; Score 19.2; DB 23; Length 1921;
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 4 atctgcagccatggcgtaggccccc 27
 ||||| ||||| || |||||
 Db 988 atctgcagccatggcgtaggccccc 1011

RESULT 6
 AAS45000
 ID AAS45000 standard; cDNA: 2390 BP.
 AC AAS45000;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE cDNA encoding novel human secretory protein, Seq ID No 81.
 XX

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen; ss.
 XX

OS Homo sapiens.

WO200166689-A2.

13-SEP-2001.

05-MAR-2001; 2001WO-US04942.

07-MAR-2000; 2000US-0519705.

19-MAY-2000; 2000US-0574454.

17-JUN-2000; 2000US-0596193.

14-JUL-2000; 2000US-0616847.

19-SEP-2000; 2000US-0655363.

20-OCT-2000; 2000US-0693267.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

WPI: 2001-589934/66.

P-PSDB; AAU28100.

Novel polypeptides and nucleic acids obtained from cDNA libraries
 prepared from various human tissues, for diagnosis and treatment of
 cancer, neurological, inflammatory, and autoimmune disorders -

Claim 1; SEQ ID No 81; 107pp; English.

The invention relates to novel isolated human secreted polypeptides (I)
 and polynucleotides (II). (I) and (II) are useful for treating
 inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 involved in increasing haematopoiesis, stem cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAS44920-AAS45295 represent novel human secreted protein
 CC coding sequences of the invention.
 XX

SQ Sequence 2390 BP; 503 A; 710 C; 723 G; 454 T; 0 other;

Query Match 66.2%; Score 19.2; DB 22; Length 2390;
 Best Local Similarity 87.5%; Pred. No. 1.1e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 4 atctgcagccatggcgtaggccccc 27

||||| ||||| || |||||

Db 1958 atctgcagccatggcgtaggccccc 1981

RESULT 7
 AAS77493/C

ID AAS77493 standard; cDNA: 3023 BP.

AC AAS77493;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #13297.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI: 2001-639362/73.

P-PSDB; ABG13306.

New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID NO 13297; 103pp; English.

XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 3023 BP; 1083 A; 698 C; 631 G; 611 T; 0 other;

Query Match 66.2%; Score 19.2; DB 23; Length 3023;

Best Local Similarity 87.5%; Pred. No. 1.1e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 tctgcagccatgctaggtgcccgcg 28
|||||
Db 1362 TCTGCAGCCATGCTAGGTGGCCG 1339

RESULT 8

AAS80007/C

ID AAS80007 standard; cDNA; 3684 BP.

XX
XX AAS80007;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #15811.

XX Human; chromosome mapping; gene mapping; gene therapy; forensics;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX
XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG15820.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID NO 15811; 103pp; English.

XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 3684 BP; 1391 A; 856 C; 712 G; 725 T; 0 other;

Query Match 66.2%; Score 19.2; DB 23; Length 3684;

Best Local Similarity 87.5%; Pred. No. 1.2e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 tctgcagccatgctaggtgcccgcg 28
|||||
Db 2628 TCTGCAGCCATGCTAGGTGGCCG 2605

RESULT 9

AAL36719

ID AAL36719 standard; DNA; 934 BP.

XX
XX AAL36719;

XX 08-JAN-2002 (first entry)

XX Human musculoskeletal system related polynucleotide SEQ ID NO 3084.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.

XX Homo sapiens.

XX WO200155367-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01338.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231988.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Example 2; SEQ ID NO 3084; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, CC

CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 934 BP; 213 A; 258 C; 241 G; 222 T; 0 other;

Query Match 65.5%; Score 19; DB 22; Length 934;
Best Local Similarity 81.5%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 gatattgcagccatggttaggcctcc 27
| ||||| | ||||| ||||| |||
Db 66 gctattcttgcctgcatgagctccc 92

RESULT 10
AAL36720
ID AAL36720 standard; DNA; 934 BP.

XX AAL36720;

XX 08-JAN-2002 (first entry)

DE Human musculoskeletal system related polynucleotide SEQ ID NO 3085.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.

XX Homo sapiens.

XX WO200155367-A1.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01338.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Example 2: SEQ ID NO 3085; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the

CC Printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 934 BP; 212 A; 258 C; 242 G; 222 T; 0 Other;
 Query Match 65.5%; Score 19; DB 22; Length 934;
 Best Local Similarity 81.5%; Pred. NO. 1.2e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 gatattgcagccatgctaggctcccc 27
 Db 66 gctattcttcgcatgtaggtctccc 92
 RESULT 11
 AAA37841
 ID AAA37841 standard; DNA; 1173 BP.
 XX
 AC AAA37841;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE Mouse angiotensin related-2 coding sequence.
 XX
 KW AR-2 fibrinogen-like domain; angiotensin related-2; mouse; ischaemia;
 KW diabetes; tumour angiogenesis; neoplastic disease; atherosclerosis;
 KW thromboembolic disease; inflammatory disease; wound healing;
 KW vascularisation; therapy; diagnosis; ds.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1173
 FT /tag= a
 FT /product= angiotensin-related-2
 XX
 PN WO200052167-A1.
 XX
 PD 08-SEP-2000.
 XX
 PF 10-FEB-2000; 2000WO-US03381.
 XX
 PR 02-MAR-1999; 99US-0122499.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Jones PF, Valenzuela DM;
 XX
 XX WPI; 2000-579286/54.
 DR P-PSDB; AAY90401.
 XX
 PT Novel nucleic acid molecules encoding fibrinogen-like domain of
 PT Angiotensin Related-2 factor useful for treating ischaemia, diabetes,
 PT and for wound healing -
 XX
 XX Example 5; Fig 5; 55pp; English.
 XX
 XX This sequence encodes mouse angiotensin related-2. It was used
 XX in the isolation of the angiotensin related-2 (AR-2) fibrinogen-like
 XX domain (FD) of the invention. AR-2 is useful for treating and diagnosing
 XX ischaemia, diabetes, tumour angiogenesis, neoplastic diseases,
 XX thromboembolic diseases, atherosclerosis, inflammatory diseases, and for
 XX wound healing. The DNA is useful for developing ligands, screening
 XX agonists and antagonists of AR-2, and as a therapeutic for treating
 XX disorders involving cells, tissues or organs expressing AR-2 receptor.
 XX AR-2 is useful to promote the growth, survival, migration, stabilisation
 XX or destabilisation, and/or differentiation of cells expressing AR-2
 XX receptor. AR-2 is also useful in assay systems to identify agonists and
 XX antagonists of AR-2 receptor. AR-2 is also useful for inducing or
 XX preventing vascularisation in diseases or disorders where such function
 XX is indicated, for delivering toxins to a receptor bearing cells and as
 XX diagnostic reagents for detecting the disease by tissue staining or

CC whole body imaging.
 XX Sequence 1173 BP; 302 A; 334 C; 316 G; 221 T; 0 other;
 SQ

Query Match 65.5%; Score 19; DB 21; Length 1173;
 Best Local Similarity 81.5%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 tatctgcagccatgctagggcccccgc 29
 || ||||| |||| ||||| ||
 Db 311 tacctgcagccagcctagtcgccccagc 337

RESULT 12
 AAS22675
 ID AAS22675 standard; cDNA; 1223 BP.
 XX AC AAS22675;
 XX DT 24-OCT-2001 (first entry)
 XX Human cDNA encoding a novel human protein #241.
 XX DE
 XX Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytosolic; neuroprotective; vulnerable; nootropic;
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.
 XX OS Homo sapiens.
 XX KW WO200155437-A2.
 XX PN 25-JAN-2001; 2001WO-US02623.
 XX PD 02-AUG-2001.
 XX PP 25-JAN-2001; 2001WO-US02623.
 XX PR 25-JAN-2000; 2000US-0491404.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-451939/48.
 XX DR P-PSDB; AAU14370.
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 XX Claim 1; Page 670-671; 894pp; English.
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,

CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence encodes a protein of the invention.
 XX Sequence 1223 BP; 226 A; 357 C; 333 G; 307 T; 0 other;
 SQ

Query Match 65.5%; Score 19; DB 22; Length 1223;
 Best Local Similarity 81.5%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatctgcagccatgctagggccccc 27
 || || ||||| ||||| |||||
 Db 1145 gaaacatcagccatgctagggccccc 1171

RESULT 13
 AAS22439
 ID AAS22439 standard; cDNA; 1260 BP.
 XX AC AAS22439;
 XX DT 24-OCT-2001 (first entry)
 XX Human cDNA encoding a novel human protein #5.
 XX DE
 XX Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytosolic; neuroprotective; vulnerable; nootropic;
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.
 XX OS Homo sapiens.
 XX KW WO200155437-A2.
 XX PN 02-AUG-2001.
 XX PD 25-JAN-2001; 2001WO-US02623.
 XX PP 25-JAN-2000; 2000US-0491404.
 XX PR (HYSE-) HYSEQ INC.
 XX PA Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-451939/48.
 XX DR P-PSDB; AAU14134.
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX
 XX Claim 1; Page 172-173; 894pp; English.
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.

Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, regenerating bone, cartilage, tendon, ligament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraceptive, treating osteoporosis and osteoarthritis, anaemia, Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral sclerosis, stroke, immune deficiencies resulting from bacterial, viral or fungal infection or from autoimmunity, cancer, allergy, asthma, graft-versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory diseases, nervous system disorders, and infection. The present sequence encodes a protein of the invention.

Sequence 1260 BP: 231 A; 371 C; 343 G; 315 T; 0 other;

```
Query Match      65.5%; Score 19; DB 22; Length 1260;
Best Local Similarity 81.5%; Pred. No. 1.3e+03;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

Qy 1 gatatctgcagccatggctaggcccc 27
|||
Db 1182 gaaacatcagccatgccaggcccc 1208

RESULT	14
AAH08668/c	
ID	AAH08668 standard; cDNA; 702 BP.
XX	
XX	
AC	AAH08668;
XX	
XX	
DT	26-JUN-2001 (first entry)
XX	
XX	
DE	Human cDNA clone (5'-primer) SEQ ID NO:5503.
XX	
XX	
XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs -
 Claim 1; SEQ ID 5503; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-*or* primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a

polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AA895893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 702 BP: 151 A; 169 C; 243 G; 136 T; 3 other;

Query Match	64.1%	Score 18.6;	DB 22;	Length 702;
Best Local Similarity	84.0%;	Pred. No. 1.8e+02;		
Matches 21;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

Qy 5 tctgcagccatggctaggccccgc 29
||| | ||||| ||| ||| |||
Db 208 TCTCCCGCCATGGCTCGGCGCCGC 184

RESULT	15
AAH06512/C	
ID	AAH06512 standard; cDNA; 703 BP.
XX	
XX	
XX	AAH06512;
XX	
XX	
XX	26-JUN-2001 (first entry)
XX	
XX	
DE	Human cDNA clone (5'-primer) SEQ ID NO:3347.

Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

Claim 1; SEQ ID 3347; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.

SQ Sequence 703 BP; 164 A; 163 C; 232 G; 141 T; 3 other;

Query Match 64.1%; Score 18.6; DB 22; Length 703;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 tctgcagccatggctaggcccccgc 29
||| ||||| ||||| |||||
Db 150 TCTCCGCCCATGGCTCGGCCGCGC 126

Search completed: August 31, 2002, 18:17:15
Job time: 16437 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:38:17 ; Search time 6064.26 Seconds
(without alignments)
64.544 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattcgagccatgtaggccccgcg 29

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:

1: em_esta:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	72.4	901	12	CNS03E19	AL240426 Tetraodon
2	20.8	71.7	365	10	BC697507	BG697507 602660715
3	20.2	69.7	687	12	AG125785	AG125785 Pan trogl
4	20	69.0	421	12	AQ458858	AQ458858 HS_5050.A
5	20	69.0	1700	10	BG111410	BG111410 602281779
6	19.6	67.6	383	9	AW346670	AW346670 29219 MAR
7	19.6	67.6	534	9	AW657502	AW657502 110747 MA
8	19.6	67.6	563	10	BC383058	BC383058 300957 MA
9	19.4	66.9	461	10	BE356665	BE356665 DGI_12.A1
10	19.4	66.9	574	9	AW066213	AW066213 687008A06
11	19.4	66.9	1249	9	AW0690022	AW0690022 602186312
12	19.2	66.2	240	10	BF516864	BF516864 NXCI_004
13	19.2	66.2	306	9	AW405340	AW405340 UI-HF-BL0
14	19.2	66.2	315	9	AW736930	AW736930 NXNV_081
15	19.2	66.2	326	10	BE761821	BE761821 NXCI_070
16	19.2	66.2	330	9	AW056925	AW056925 ST57H10.P
17	19.2	66.2	330	10	BE582369	BE582369 NXCI_032

18	19.2	66.2	348	10	BF220344	BF220344 NXCI_142
19	19.2	66.2	400	10	BF777295	BF777295 NXSI_067
20	19.2	66.2	423	10	H02247	H02247 Y735b01.s1
21	19.2	66.2	443	10	BE761860	BE761860 NXCI_070
22	19.2	66.2	452	10	BE656912	BE656912 NXCI_057
23	19.2	66.2	518	12	AZ827981	AZ827981 2M0104815
24	19.2	66.2	519	10	BG040220	BG040220 NXSI_109
25	19.2	66.2	534	10	BF610423	BF610423 NXSI_038
26	19.2	66.2	538	10	BF610478	BF610478 NXSI_058
27	19.2	66.2	548	9	AW290141	AW290141 NXNV012B0
28	19.2	66.2	591	9	AW064983	AW064983 ST38C01.P
29	19.2	66.2	618	11	AK017620	AK017620 Mus muscu
30	19.2	66.2	983	9	AL572399	AL572399 AL572399
31	19.2	66.2	1100	12	BG286946	BG286946 602382830
32	19	65.5	388	10	AZ995875	AZ995875 2M0281C20
33	19	65.5	452	12	AQ495929	AQ495929 HS_5056.A
34	19	65.5	490	10	BJ202596	BJ202596 BJ202596
35	19	65.5	554	10	BJ205326	BJ205326 BJ205326
36	19	65.5	561	10	BJ202480	BJ202480 BJ202480
37	19	65.5	563	10	BJ192283	BJ192283 BJ192283
38	19	65.5	573	10	BJ156941	BJ156941 BJ156941
39	19	65.5	579	10	BJ200115	BJ200115 BJ200115
40	19	65.5	594	10	BF282124	BF282124 EST446715
41	19	65.5	673	12	AZ359142	AZ359142 1M0101C19
42	19	65.5	1125	10	BG025522	BG025522 60275058
43	19	65.5	2306	11	AK011976	AK011976 Mus muscu
44	18.8	64.8	150	10	BF604299	BF604299 270077 MA
45	18.8	64.8	227	9	AA556708	AA556708 563 Loblo

ALIGNMENTS

RESULT 1

CNS03E19
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
020612 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL240426
AL240426.1 GI:7961195
GSS: genome survey sequence.
Tetraodon nigroviridis.
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 901)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 901)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 901)
Direct Submission
Genoscope.
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1..901
/organism="Tetraodon nigroviridis"

```

/db_xref="taxon:99883"
/clone="020G12"
/clone_lib="G"
/note="Genoscope sequence ID : COBG020BD06SP1-end :
PUC-ori"
BASE COUNT 254 a 239 c 228 g 173 t 7 others
ORIGIN

Query Match 72.4%; Score 21; DB 12; Length 901;
Best Local Similarity 82.8%; Pred. No. 4.5e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatattgcagccatgctagggccccgc 29
|||||
Db 154 GATCTCTGCAGCCATGCTCGGCTCCGCC 182

RESULT 2
BG697507/c
LOCUS
DEFINITION
602660715F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4804011 5',
mRNA sequence.
ACCESSION BG697507
VERSION BG697507
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 365)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@rsr@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10700 row: m column: 04
High quality sequence start: 9
High quality sequence stop: 365.
Location/Qualifiers
1..365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4804011"
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 62 a 117 c 121 g 65 t
ORIGIN

Query Match 71.7%; Score 20.8; DB 10; Length 365;
Best Local Similarity 91.7%; Pred. No. 4.4e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ctgcagccatgctagggccccgc 29
|||||
Db 54 CTCAGCCATGGCCAGCCCCAGC 31

RESULT 3
AG125785/c
LOCUS

```

```

DEFINITION
ACCESSION AG125785
VERSION AG125785.1 GI:16654950
KEYWORDS GSS; GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-136C09.F.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (sites)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 687)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chihirobes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..687
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-136C09.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 190 a 142 c 140 g 213 t 2 others
ORIGIN

Query Match 69.7%; Score 20.2; DB 12; Length 687;
Best Local Similarity 88.0%; Pred. No. 8.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gatattgcagccatgctagggccc 25
|||||
Db 331 GATATCTGAAGCCATGCGCAAGACCC 307

RESULT 4
AQ458858
LOCUS
DEFINITION
HS_5050_AL_A10_SP6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=626 Col=19 Row=A, DNA sequence.
ACCESSION AQ458858
VERSION AQ458858.1 GI:4637498
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 421)
Mahairas, G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
MEDLINE

```


COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 626 row: A column: 19
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 421.

FEATURES

source

1. .421
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-626 Col-19 Row-A"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

BASE COUNT

136 a 70 c 69 g 145 t 1 others

ORIGIN

Query Match

Best Local Similarity 69.0%; Score 20; DB 12; Length 421;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 atatctgcagccatggttaggccccgcg 29

Db 251 ATATCTCGAGCATTTTGGACCCAGC 278

RESULT 5

BG111410/c 1700 bp mRNA linear EST 30-JAN-2001
LOCUS 602281779F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369363 5',
DEFINITION mRNA sequence.

ACCESSION

BG111410

VERSION

BG111410.1 GI:12604832

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1700)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLNL0024 Row: n column: 20
High quality sequence stop: 290.

FEATURES

source

1. .1700
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:4369363"
/clone_lib="NIH_MGC_86"

/tissue_type="osteosarcoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.533 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 582 a 552 c 451 g 115 t

ORIGIN

Query Match

Best Local Similarity 69.0%; Score 20; DB 10; Length 1700;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gatattctgcagccatggttaggccccgcg 28

Db 594 GGTATCTTCGCCCATGGCTAGGCTCGCG 567

RESULT 6

AW346670

LOCUS

AW346670 383 bp mRNA linear EST 09-JUL-2000

DEFINITION

29219 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.

ACCESSION

AW346670

VERSION

AW346670.1 GI:6844380

KEYWORDS

EST.

SOURCE

plg.

ORGANISM

Sus scrofa

REFERENCE

1 (bases 1 to 383)

AUTHORS

Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W. and Reele, J.W.

TITLE

Design and use of two pooled tissue normalized cDNA libraries for

JOURNAL

EST discovery in swine

COMMENT

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 20

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG

Plate: 13 row: P column: 6

Seq primer: ATTTAGTGACACTAG.

FEATURES

source

1. .383

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone_lib="MARC 1P1G"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from day 11, 13, 15, 20,

and 30 embryos."

BASE COUNT 89 a 141 c 103 g 50 t

ORIGIN

Query Match

Best Local Similarity 67.6%; Score 19.6; DB 9; Length 383;

Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 atatctgcagccatggttaggccccgcg 27

DB 346 ATACCTGCAGCCATGGCAGTGGCCCC 371
|||||

RESULT 7

AW657502 554 bp mRNA linear EST 09-JUL-2000
LOCUS 110747 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION AW657502
ACCESSION AW657502
VERSION AW657502.1 GI:7423328
KEYWORDS EST.
SOURCE pig.

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 554)
Fahrenkrug, S.C., Fraking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 38 row: 1 column: 2
Seq primer: ATTTAGTGACACTATAG.

FEATURES
source
1..554
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
Location/Qualifiers

BASE COUNT

117 a 198 c 154 g 85 t

Query Match 67.6%; Score 19.6; DB 9; Length 554;
Best Local Similarity 84.6%; Pred. No. 1.4e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggcccc 27
|||||

DB 419 ATACCTGCAGCCATGGCAGTGGCCCC 444

RESULT 8

BG383058 563 bp mRNA linear EST 12-MAR-2001
LOCUS 300957 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG383058
ACCESSION BG383058
VERSION BG383058.1 GI:13307530
KEYWORDS EST.
SOURCE pig.

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 563)
Fahrenkrug, S.C., Fraking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.

TITLE
JOURNAL
COMMENT

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 6 row: 1 column: 20
Seq primer: ATTTAGTGACACTATAG.

FEATURES
source
1..563
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
Location/Qualifiers

BASE COUNT 165 a 148 c 129 g 120 t 1 others
ORIGIN

Query Match 67.6%; Score 19.6; DB 10; Length 563;
Best Local Similarity 84.6%; Pred. No. 1.4e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggcccc 27
|||||

DB 359 ATGAATGCAGCCAGGCTAGGCCCC 334

RESULT 9

BE356665/c 461 bp mRNA linear EST 20-JUL-2000
LOCUS DGI_12.A12.b1.A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
DEFINITION BE356665
ACCESSION BE356665
VERSION BE356665.1 GI:9298222
KEYWORDS EST.
SOURCE sorghum.

ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 461)
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt
L.H.

REFERENCE An EST database from Sorghum: dark-grown seedlings
AUTHORS Unpublished (2000)
JOURNAL Contact: Cordonnier-Pratt MM
COMMENT Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 451
POLYA-No. Location/Qualifiers

FEATURES

[illegible]

```

SOURCE      loblolly pine.
ORGANISM     Pinus taeda

REFERENCE    Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
1 (bases 1 to 240)
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS      Sederoff, R.
TITLE        Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL      Unpublished (2000)
COMMENT      Contact: Johnson, Arthur
            North Carolina State University
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ajohnson@unity.ncsu.edu
            Seq primer: T3.

FEATURES     Location/Qualifiers
source       1..240
            /organism="Pinus taeda"
            /strain="Coastal plain loblolly pine from North Carolina"
            /db_xref="taxon:3352"
            /clone="NXSI_004_D07"
            /clone_lib="NXSI (Nsf Xylem Side wood Inclined)"
            /tissue_type="Xylem"
            /cell_type="Side"
            /dev_stage="Juvenile"
            /lab_host="XLI-Blue"
            /note="vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
            ; The library is from early (spring) wood, taken from
            three six-year old trees (three different genotypes), in
            the juvenile phase. These trees were induced to form side
            wood by bending to a 45 degree angle and tying them to the
            ground. Differentiating xylem was harvested from the sides
            of the inclined stems, and a mixture of all three
            genotypes was used for the library. oligo-dt primed cDNA
            was directionally cloned into the EcoRI-XhoI Bluescript SK
            vector arms. NOTE: The sequences contain a 'cDNA adapter'
            between the EcoRI site and the start of the EST. The
            adapter sequence is 'AATTGGGCACGAG'."

BASE COUNT   55 a 70 c 63 g 44 t 8 others
ORIGIN

Query Match 66.2%; Score 19.2; DB 10; Length 240;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 tctgcagccatggttaggccccg 28
    ||| ||||| || |||||
Db 27 TCTGAAGCCATGGCAAGACCCCG 50

RESULT 13
AW405340
LOCUS      UT-HF-BLO-ack-f-06-0-UI.r1 NTH_MGC_37 Homo sapiens cDNA clone
DEFINITION IMAGE:3059459 5', mRNA sequence.
ACCESSION  AW405340
VERSION     AW405340.1 GI:6924397
KEYWORDS   EST.
SOURCE      Human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 306)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabs@emil.nih.gov
            Eco RI site shown at the beginning of the sequence.
            Tissue procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: M.B. Soares Lab
            cDNA Library Arrayed by: M.B. Soares Lab
            DNA Sequencing by: M.B. Soares Lab

```

```

Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.

FEATURES     Location/Qualifiers
source       1..306
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3059459"
            /clone_lib="NIH_MGC_37"
            /tissue_type="lymph"
            /cell_type="germinal center B cells"
            /cell_line="MGC85"
            /lab_host="DH10B (LTI)"
            /note="vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
            Constructed from size fractionated cytoplasmic mRNA
            (1.5-2.5kb). Directionally cloned. Cells provided by Louis
            M. Staudt, Ph.D. Library preparation by Maria de Fatima
            Ronaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT   61 a 117 c 88 g 40 t
ORIGIN

Query Match 66.2%; Score 19.2; DB 9; Length 306;
Best Local Similarity 87.5%; Pred. No. 1.7e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 atctgcagccatggttaggcccc 27
    ||| ||||| || |||||
Db 272 ATCTGCACCATGGCATGGACCCC 295

RESULT 14
AW736930
LOCUS      NXNV_081_F12_F Nsf Xylem Normal wood Vertical Pinus taeda cDNA
DEFINITION clone NXNV_081_F12 5', mRNA sequence.
ACCESSION  AW736930
VERSION     AW736930.1 GI:7644794
KEYWORDS   EST.
SOURCE      loblolly pine.
ORGANISM    Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
            1 (bases 1 to 315)
            Sederoff, R.
            Molecular Basis of Wood Formation in the Pine Megagenome
            Unpublished (2000)
            Contact: Johnson, Arthur
            North Carolina State University
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ajohnson@unity.ncsu.edu
            Seq primer: T3.

FEATURES     Location/Qualifiers
source       1..315
            /organism="Pinus taeda"
            /db_xref="taxon:3352"
            /clone="NXNV_081_F12"
            /clone_lib="Nsf Xylem Normal wood Vertical"
            /note="vector: Bluescript SK; Site_1: Eco RI; The
            sequences contain a 'cDNA adapter' between the EcoRI site
            and the start of the EST. The adapter sequence is
            'AATTGGGCACGAG'."

BASE COUNT   79 a 87 c 68 g 70 t 11 others
ORIGIN

Query Match 66.2%; Score 19.2; DB 9; Length 315;
Best Local Similarity 87.5%; Pred. No. 1.7e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 tctgcagccatggttaggccccg 28

```

||||| ||||||||| || |||||||
Db 155 TCTGAGCCATGGCAGACCCCG 178

RESULT 15

BE761821
LOCUS NXCI_070_B11_F NXCI (Nsf Xylem Compression wood Inclined) Pinus. 326 bp mRNA linear EST 19-SEP-2000
DEFINITION taeda cDNA clone NXCI_070_B11 5', mRNA sequence.

ACCESSION BE761821

VERSION BE761821.1 GI:10191745

KEYWORDS EST.

SOURCE loblolly pine.

ORGANISM Pinus taeda

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.

AUTHORS 1 (bases 1 to 326)

TITLE Sederoff, R.

JOURNAL Molecular Basis of Wood Formation in the Pine Megagenome

COMMENT Unpublished (2000)

Contact: Johnson, Arthur

North Carolina State University

Tel: 919 515 7800

Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu

Seq primer: T3.

Location/Qualifiers

1..326
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXCI_070_B11"
/clone_lib="NXCI (Nsf Xylem Compression wood Inclined)"
/tissue_type="Xylem"
/cell_type="Compression"
/dev_stage="Juvenile"
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form
compression wood by bending to a 45 degree angle and tying
them to the ground. Differentiating xylem was harvested
from the bottoms of the inclined stems, and a mixture of
all three genotypes was used for the library. Oligo-dr
primed cDNA was directionally cloned into the EcoRI-XhoI
Bluescript SK vector arms. NOTE: The sequences contain a
'cDNA adapter' between the EcoRI site and the start of the
EST. The adapter sequence is 'AATTCGGCACGAG'."

BASE COUNT 70 a 96 c 92 g 56 t 12 others

ORIGIN

Query Match 66.2%; Score 19.2; DB 10; Length 326;
Best Local Similarity 87.5%; Pred No. 1.7e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 5 tctgcagccatggctagggccccc 28

||||| ||||||||| || |||||||

Db 97 TCTGAGCCATGGCAGACCCCG 120

Search completed: August 31, 2002, 15:24:33
Job time: 6376 sec

THIS PAGE BLANK (ver. 1.0)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:40:38 ; Search time 153.1 Seconds
(without alignments)
46.528 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattgcagccatggctaggcccccgc 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18.2	62.8	3107	1	US-07-846-181-4
C 2	18.2	62.8	3107	1	US-07-845-989-4
C 3	17.8	61.4	1774	1	US-08-377-292-3
C 4	17.8	61.4	4353	2	US-08-365-486A-18
C 5	17.8	61.4	4353	4	US-08-880-342-18
C 6	17.8	61.4	4780	2	US-08-365-486A-20
C 7	17.8	61.4	4780	3	US-09-123-708-3
C 8	17.8	61.4	4780	3	US-09-123-624-3
C 9	17.8	61.4	4780	3	US-08-880-342-20
C 10	17.6	60.7	652	4	US-08-998-416-962
C 11	17.2	59.3	31	2	US-08-890-980-80
C 12	17.2	59.3	31	2	US-08-890-980-82
C 13	17.2	59.3	31	3	US-08-890-979-71
C 14	17.2	59.3	31	3	US-08-890-979-73
C 15	17.2	59.3	31	3	US-09-032-894-80
C 16	17.2	59.3	31	3	US-09-032-894-82
C 17	17.2	59.3	31	4	US-09-031-626-80
C 18	17.2	59.3	31	4	US-09-031-626-82
C 19	17.2	59.3	160	2	US-08-890-980-66
C 20	17.2	59.3	160	3	US-08-890-979-65
C 21	17.2	59.3	160	3	US-09-032-894-66
C 22	17.2	59.3	160	4	US-09-031-626-66
C 23	17.2	59.3	730	4	US-09-342-084-5
C 24	16.8	57.9	800	4	US-08-998-416-390
C 25	16.8	57.9	5975	1	US-08-920-812-23
C 26	16.8	57.9	5975	1	US-08-920-827-23
C 27	16.8	57.9	5975	1	US-08-921-177-23

28 16.8 57.9 5975 1 US-08-362-577C-23 Sequence 23, Appl
29 16.8 57.9 5975 2 US-08-920-828-23 Sequence 23, Appl
C 30 16.6 57.2 4403765 4 US-09-103-840A-2 Sequence 2, Appl
C 31 16.6 57.2 4411529 4 US-09-103-840A-1 Sequence 1, Appl
C 32 16.4 56.6 1704 1 US-08-007-282B-1 Sequence 1, Appl
C 33 16.2 55.9 376 1 US-08-785-530-2 Sequence 2, Appl
C 34 16.2 55.9 376 2 US-09-123-850-2 Sequence 2, Appl
C 35 16.2 55.9 515 4 US-09-060-756-656 Sequence 656, App
C 36 16.2 55.9 638 4 US-08-858-207A-142 Sequence 142, App
C 37 16.2 55.9 779 4 US-08-998-416-451 Sequence 451, App
C 38 16.2 55.9 1685 1 US-07-708-038-1 Sequence 1, Appl
C 39 16.2 55.9 1685 1 US-08-127-995-1 Sequence 1, Appl
C 40 16.2 55.9 5057 2 US-08-365-486A-12 Sequence 12, Appl
C 41 16.2 55.9 5057 4 US-08-880-342-12 Sequence 12, Appl
C 42 16.2 55.9 5108 1 US-07-642-002-1 Sequence 1, Appl
C 43 16.2 55.9 9143 2 US-08-639-857-32 Sequence 32, Appl
C 44 16.2 55.9 4403765 4 US-09-103-840A-2 Sequence 2, Appl
C 45 16.2 55.9 4411529 4 US-09-103-840A-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-846-181-4/c
; Sequence 4, Application US/07846181
; Patent No. 5360732
; GENERAL INFORMATION:
; APPLICANT: BERKA, RANDY M
; APPLICANT: FOWLER, TIMOTHY
; APPLICANT: REY, MICHAEL W
; TITLE OF INVENTION: PRODUCTION OF ASPERGILLUS NIGER
; TITLE OF INVENTION: CATALASE-R
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENECOR INTERNATIONAL, INC.
; STREET: 180 KIMBALL WAY
; CITY: SOUTH SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/846,181
; FILING DATE: 19920304
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HORN Ms, MARGARET A
; REGISTRATION NUMBER: 33401
; REFERENCE/DOCKET NUMBER: GC204-US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-742-7536
; TELEFAX: 415-742-7217
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3107 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(327, 620, 683...907, 969...1385, 1440...1604,
; LOCATION: 1654...2745)
US-07-846-181-4

Query Match 62.8%; Score 18.2; DB 1; Length 3107;
Best Local Similarity 87.0%; Pred. No. 51;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggcc 24
||||| ||| ||||| |||||
Db 242 ATATCGCATCCATGGCGAGGCC 220

RESULT 2
US-07-845-989-4/c
; Sequence 4, Application US/07845989
; Patent No. 5360901
; GENERAL INFORMATION:
; APPLICANT: BERKA, RANDY M
; APPLICANT: FOWLER, TIMOTHY
; APPLICANT: REY, MICHAEL W
; TITLE OF INVENTION: PRODUCTION OF ASPERGILLUS NIGER
; TITLE OF INVENTION: CATALASE-R
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENECOR INTERNATIONAL, INC.
; STREET: 180 KIMBALL WAY
; CITY: SOUTH SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/845,989
; FILING DATE: 19920304
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HORN MS, MARGARET A
; REGISTRATION NUMBER: 33401
; REFERENCE/DOCKET NUMBER: GC208-US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-742-7536
; TELEFAX: 415-742-7217
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3107 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(327..620, 583..907, 969..1385, 1440..1604,
; LOCATION: 1654..2745)
US-07-845-989-4

Query Match 62.88; Score 18.2; DB 1; Length 3107;
Best Local Similarity 87.08; Pred. No. 51;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 atatctgcagccatggctaggcc 24
||||| ||| ||||| |||||
Db 242 ATATCGCATCCATGGCGAGGCC 220

RESULT 3
US-08-377-292-3/c
; Sequence 3, Application US/08377292
; Patent No. 5693615
; GENERAL INFORMATION:
; APPLICANT: STONE, ROGER L.
; TITLE OF INVENTION: THERAPEUTIC FORMULAS FOR OSTEOINDUCTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road

; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 45239-8707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,292
; FILING DATE: 23-JAN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,435
; FILING DATE:
; APPLICATION NUMBER: US/08/117,367
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corstanje, Brahm J.
; REGISTRATION NUMBER: 34,804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-245-2858
; TELEFAX: 513-741-3012
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-377-292-3

Query Match 61.48; Score 17.8; DB 1; Length 1774;
Best Local Similarity 90.58; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 9 cagccatggctaggcccccgc 29
||||| ||||| |||||
Db 327 CAGCCATGGCTAGGTACCCGC 307

RESULT 4
US-08-365-486A-18/c
; Sequence 18, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,486A
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4353 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human NOS-1 gene, Fujisawa, et al,
INDIVIDUAL ISOLATE: J. Neurochem 63:140 1994
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4305
US-08-365-486A-18

Query Match 61.4%; Score 17.8; DB 2; Length 4353;
Best Local Similarity 75.9%; Pred. No. 77;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gatattcgagccatggtaggcccccgc 29
|| ||| ||||| ||||| || |||||
Db 4089 GACATCAGCAGCCATGTCAGTCCCCAC 4061

RESULT 5
US-08-880-342-18/c
Sequence 18, Application US/08880342
Patent No. 6218179
GENERAL INFORMATION:
APPLICANT: Webster, Keith A.
APPLICANT: Bishopric, Nanette H.
APPLICANT: Murphy, Brian
APPLICANT: Laderoute, Keith R.
APPLICANT: Green, Christopher J.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
TITLE OF INVENTION: Therapeutic Constructs
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,342
FILING DATE: 23-JUN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB95/00996
FILING DATE: 13-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/365,486
FILING DATE: 23-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 4353 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human NOS-1 gene, Fujisawa, et al,
INDIVIDUAL ISOLATE: J. Neurochem 63:140 1994
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4305
US-08-880-342-18

Query Match 61.4%; Score 17.8; DB 4; Length 4353;
Best Local Similarity 75.9%; Pred. No. 77;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gatattcgagccatggtaggcccccgc 29
|| ||| ||||| ||||| || |||||
Db 4089 GACATCAGCAGCCATGTCAGTCCCCAC 4061

RESULT 6
US-08-365-486A-20/c
Sequence 20, Application US/08365486A
Patent No. 5834306
GENERAL INFORMATION:
APPLICANT: Webster, Keith A.
APPLICANT: Bishopric, Nanette H.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
TITLE OF INVENTION: Therapeutic Constructs
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,486A
FILING DATE: 23-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 4780 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human NOS-SN gene, Nakane, et al,
INDIVIDUAL ISOLATE: FEBS Lett 316:175 (1993)
FEATURE:
NAME/KEY: CDS

LOCATION: 431..4732
US-08-365-486A-20

Query Match 61.4%; Score 17.8; DB 2; Length 4780;
Best Local Similarity 75.9%; Pred. NO. 78;
Matches 22; Conservative 0; Mismatches 7; Indels 0

Qy 1 gatatctgcagccatggctaggccccgc 29
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4516 GACATCAGCAGCCATGGTGACGTCCCCAC 4488

RESULT 7
US-09-123-708-3/c
; Sequence 3, Application US/09123708

```

: GENERAL INFORMATION:
: APPLICANT: SCHRADER, Juergen
: APPLICANT: GORDECKE, Axel
: TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
: TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS

```

```

; FILE NAME: D:\BRENCH\2006-07-28\1005
; CURRENT APPLICATION NUMBER: US/09/133,708
; CURRENT FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/553,503
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: P4411402.8
; EARLIER FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0

```

; LENGTH: 4780
 ; TYPE: DNA
 ; ORGANISM: Cytomegalovirus
 US-09-123-708-3

Query Match 61.4%; Score 17.8; DB 3; Length 4780;
Best Local Similarity 75.9%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches 7; Indels 0

Qy · 1 gatatctgcagccatggctaggccccgc 29
||| ||| ||| ||| ||| ||| ||| ||| |||
Dd 4516 GACATCAGCAGCCCATGGTGACGTCCCCAC 4488

RESULT 8
US-09-123-624-3/c
; Sequence 3, Application US/091233624

```

/ GENETIC SOURCE NO.: 0497250
/ GENERAL INFORMATION:
/ APPLICANT: SCHRAEDER, Jürgen
/ APPLICANT: GODECKE, Axel
/ TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
/ TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS

```

```

, FILE REFERENCE: 31109-2004
, CURRENT APPLICATION NUMBER: US/09/123,624
, CURRENT FILING DATE: 1998-07-28
, PRIOR APPLICATION NUMBER: 08/553,503
, PRIOR FILING DATE: 1996-03-01
, PRIOR APPLICATION NUMBER: 441402.8
, PRIOR FILING DATE: 1994-03-31
, NUMBER OF SEQ ID NOS: 6
, SOFTWARE: Patent Ver. 2.1

```

```

; SEQ ID NO 3
; LENGTH: 4780
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-123-624-3

```

Query Match 61.48; Score 17.8; DB 3; Length 4780;

Best Local Similarity 75.9%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches

Qy 1 gatatctgcagccatggctaggcccccg 29
 || ||| ||||||||| | |||||
Dδ 4516 GACATCAGCAGCCATGGTGACGTCCCCAC 4488

RESULT 9
US-08-880-342-20/c
; Sequence 20, Application US/08880342

; GENERAL INFORMATION.
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; APPLICANT: Murphy, Brian

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

ADDRESS: Dellinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/880,342
; FILING DATE:  23-JUN-1997
;

```

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/IB95/00996
 ; FILING DATE: 13-NOV-1995
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/363,480
 ; FILING DATE: 23-DEC-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sholtz, Charles K.

TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

```

; INFORMATION FOR SEQ ID NO: 2
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 4780 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear

```

```

1 MOLECULE TYPE: cDNA to mRNA
2
3 HYPOTHETICAL: NO
4
5 ANTI-SENSE: NO
6
7 ORIGINAL SOURCE:
8
9 INDIVIDUAL ISOLATE: Human NOS-SN gene, Nakane, et al.
10
11 INDIVIDUAL ISOLATE: FEBS Lett 316:175 (1993)
12
13 FEATURE:
14
15 NAME/KEY: CDS
16
17 LOCATION: 431 4732
18

```

US-08-880-342-20	Query Match	Best Local Similarity
6		
7		

Qy 1 qatatctgcaqcccatgqctagqcccccc 29

Query Match 61.4%; Score 17.8; DB 4; Length 4780;
Best Local Similarity 75.9%; pred. No. 78;
Matches 22; Conservative 0; Mismatches 7; Indels 0

Db 4516 GACATCAGCAGCCATGGTGCAGTCCAC 4488
|| ||| ||||| ||||| || |||||

RESULT 10

US-08-998-416-962/c
; Sequence 962, Application US/08998416
; Patent No. 6239264

; GENERAL INFORMATION:

; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jürgen
; APPLICANT: Knechtle, Philipp

; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: CH 0016/97

; FILING DATE: 31-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Meigs, J. Timothy

; REGISTRATION NUMBER: 38,241

; REFERENCE/DOCKET NUMBER: PE/5-30306/A/CGC1976

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8587

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 962:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 652 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: PAG1593UP

US-08-998-416-962

Query Match 60.7%; Score 17.6; DB 4; Length 652;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 tatctcagccatggctaggccccc 26

||||| ||||| ||||| |||||

Db 87 TATCTATAGCCATCGCCAGGCCCC 64

RESULT 11

US-08-890-980-80

; Sequence 80, Application US/08890980

; Patent No. 5998141

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan L.

; TITLE OF INVENTION: SR-B1 NUCLEIC ACIDS AND USES THEREFOR

; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/890,980

; FILING DATE: 10-JUL-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold, Beth E.

; REGISTRATION NUMBER: 35,430

; REFERENCE/DOCKET NUMBER: MIA-005.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-832-1000

; TELEFAX: 617-832-7000

; INFORMATION FOR SEQ ID NO: 80:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 31 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "probe"

US-08-890-980-80

Query Match 59.3%; Score 17.2; DB 2; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 gcagccatggctaggcccccgc 29

||||| ||||| ||||| |||||

Db 4 GCAGCCATGGCCAGGCCACCC 25

RESULT 12

US-08-890-980-82/c

; Sequence 82, Application US/08890980

; Patent No. 5998141

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan L.

; TITLE OF INVENTION: SR-B1 NUCLEIC ACIDS AND USES THEREFOR

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FOLEY, HOAG & ELIOT LLP

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109-2170

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/890,980

; FILING DATE: 10-JUL-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold, Beth E.

; REGISTRATION NUMBER: 35,430

; REFERENCE/DOCKET NUMBER: MIA-005.01

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "probe"
US-08-890-980-82

Query Match 59.3%; Score 17.2; DB 2; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 gcagccatggctagcccccgc 29
|||||
Db 28 GCAGCCATGGCCAGGCCACCC 7

RESULT 13
US-08-890-979-71
Sequence 71, Application US/08890979
Patent No. 6030778
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
APPLICANT: Ordovas, Jose M.
TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS
TITLE OF INVENTION: DISORDERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/890,979
FILING DATE: 10-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-005.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "probe"
US-08-890-979-71

Query Match 59.3%; Score 17.2; DB 3; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 gcagccatggctagcccccgc 29
|||||
Db 4 GCAGCCATGGCCAGGCCACCC 25

RESULT 14
US-08-890-979-73/c
Sequence 73, Application US/08890979
Patent No. 6030778
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
APPLICANT: Ordovas, Jose M.
TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS
TITLE OF INVENTION: DISORDERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/890,979
FILING DATE: 10-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-005.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "probe"
US-08-890-979-73

Query Match 59.3%; Score 17.2; DB 3; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 gcagccatggctagcccccgc 29
|||||
Db 28 GCAGCCATGGCCAGGCCACCC 7

RESULT 15
US-09-032-894-80
Sequence 80, Application US/09032894
Patent No. 6130041
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
TITLE OF INVENTION: SR-BI NUCLEIC ACIDS AND USES THEREFOR
FILE REFERENCE: MIA-005.03
CURRENT APPLICATION NUMBER: US/09/032,894
CURRENT FILING DATE: 1998-02-27
EARLIER APPLICATION NUMBER: 08/890,980
EARLIER FILING DATE: 1997-07-10
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 80
LENGTH: 31
TYPE: DNA
ORGANISM: Human

US-09-032-894-80

Query Match 59.3%; Score 17.2; DB 3; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 gcagccatggctaggccccgc 29
| | | | | | | | | | | | | |
Db 4 gcagccatggccagcccc 25

Search completed: August 31, 2002, 15:27:36
Job time: 6418 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:03:50 ; Search time 9324.87 Seconds
(without alignments)
69.569 Million cell updates/sec

Title: US-09-810-861b-2
Perfect score: 31
Sequence: 1 cggtagctatcaggtagcgtgagcaatttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vl.*
30: em.htg.hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
------------	-------	-------	--------------	-------	-------------

1	31	100.0	31	6	AX275253	Sequence
2	31	100.0	5767	6	AX275254	Sequence
3	31	100.0	14446	6	AX275255	Sequence
4	20.4	65.8	1725	6	AX275256	Sequence
5	20.4	65.8	136257	9	AC073864	Homo sapi
6	20.4	65.8	170075	2	AC026678	Homo sapi
7	20.4	65.8	172241	2	AC024508	Homo sapi
8	20.4	65.8	178820	9	AC020552	Homo sapi
9	20.4	65.8	181230	2	AL358876	Homo sapi
10	20.4	65.8	183754	2	AC023183	Homo sapi
11	20.4	65.8	260430	2	AC024935	Homo sapi
12	20	64.5	1019	1	DS276564	Desulfoto
13	20	64.5	1409	1	AB015893	Treponema
14	20	64.5	1410	1	AB015888	Treponema
15	20	64.5	1414	1	AB015889	Treponema
16	20	64.5	1417	1	AB015887	Treponema
17	20	64.5	1421	1	AB015892	Treponema
18	20	64.5	1475	1	AF033303	Treponema
19	20	64.5	1479	1	AF033307	Treponema
20	20	64.5	1543	1	AF033306	Treponema
21	20	64.5	87600	9	AC007160	Homo sapi
22	19.8	63.9	478	1	USP287081	unculture
23	19.8	63.9	1469	1	AF023038	Treponema
24	19.8	63.9	32147	9	AL358613	Human DNA
25	19.8	63.9	348550	1	AP003596	Nostoc sp
26	19.6	63.2	73814	2	AC069345	Homo sapi
27	19.6	63.2	95180	9	AC098853	Homo sapi
28	19.6	63.2	218677	2	AC104842	Homo sapi
29	19.4	62.6	461	1	USP287095	unculture
30	19.4	62.6	91316	2	AC094771	Rattus no
31	19.4	62.6	93957	2	AC021819	Homo sapi
32	19.4	62.6	110000	2	LMFLCHR36_27	Continuation (28 o
33	19.4	62.6	110000	2	LMFLCHR36_28	Continuation (29 o
34	19.4	62.6	141180	2	AC078807	Homo sapi
35	19.4	62.6	141203	2	AC068635	Homo sapi
36	19.4	62.6	170940	2	AC069067	Homo sapi
37	19.4	62.6	180681	9	AC017016	Homo sapi
38	19.2	61.9	1471	1	BLSF016SN	X80834 Long segmen
39	19	61.3	1637	1	AF231923	AF231923 xanthomon
40	19	61.3	64414	9	AC023792	Homo sapi
41	19	61.3	145540	2	AC015962	Homo sapi
42	19	61.3	171718	2	AC022501	Homo sapi
43	19	61.3	173037	2	AC018542	Homo sapi
44	19	61.3	173487	2	AC026925	Homo sapi
45	19	61.3	174803	9	AL591501	Human DNA

ALIGNMENTS

RESULT	1					
AX275253	AX275253	31 bp	DNA	linear	PAT 29-OCT-2001	
LOCUS	Sequence 2 from Patent WO0171014.					
DEFINITION	AX275253					
ACCESSION	AX275253.1					
VERSION	AX275253.1	GI:16547673				
KEYWORDS	synthetic construct.					
SOURCE	synthetic construct.					
ORGANISM	artificial sequence.					
REFERENCE	1 (sites)					
AUTHORS	Mor,T., Soreq,H., Arntzen,C. and Mason,H.					
TITLE	Expression of recombinant human acetylcholinesterase in transgenic plants					
JOURNAL	Patent: WO 0171014-A 2 27-SEP-2001;					
	BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,					
	Tsaifir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,					
	Hugh S. (US)					
FEATURES	Location/Qualifiers					
source	1..31					
	/organism="synthetic construct"					
	/db_xref="taxon:32630"					
	/note="primer pAChE-Kpn, derived from human AChE gene and					

BASE COUNT 7 a 7 c 9 g 8 t
 ORIGIN modified to introduce a Kpn I restriction site"

Query Match 100.0%; Score 31; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.0028;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 cgggtacccatcaggtgcgctgagcaatttg 31
 |||||
 Db 1 CGGTACCTATCAGGTAGCGCTGAGCAATTG 31

RESULT 2
 AX275254/c
 LOCUS AX275254 5767 bp DNA linear PAT 29-OCT-2001
 DEFINITION Sequence 3 from Patent WO0171014.
 ACCESSION AX275254
 VERSION AX275254.1 GI:16547674
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (sites)
 AUTHORS Mor, T., Soreq, H., Arntzen, C. and Mason, H.
 TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
 JOURNAL Patent: WO 0171014-A 3 27-SEP-2001;
 BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US); Mor,
 Tsafirir (US); Soreq, Hermona (IL); Arntzen, Charles (US); Mason,
 Hugh S. (US)

FEATURES
 source
 1. 5767
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="plasmid vector pTM034"
 BASE COUNT 1390 a 1521 c 1495 g 1361 t
 ORIGIN

Query Match 100.0%; Score 31; DB 6; Length 5767;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 cgggtacccatcaggtgcgctgagcaatttg 31
 |||||
 Db 2571 CGGTACCTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
 AX275255/c
 LOCUS AX275255 14446 bp DNA linear PAT 29-OCT-2001
 DEFINITION Sequence 4 from Patent WO0171014.
 ACCESSION AX275255
 VERSION AX275255.1 GI:16547675
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (sites)
 AUTHORS Mor, T., Soreq, H., Arntzen, C. and Mason, H.
 TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
 JOURNAL Patent: WO 0171014-A 4 27-SEP-2001;
 BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US); Mor,
 Tsafirir (US); Soreq, Hermona (IL); Arntzen, Charles (US); Mason,
 Hugh S. (US)

FEATURES
 source
 1. 14446
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="plasmid vector pTM036"

BASE COUNT 3231 a 3831 c 4046 g 3042 t 296 others
 ORIGIN

Query Match 100.0%; Score 31; DB 6; Length 14446;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 cgggtacccatcaggtgcgctgagcaatttg 31
 |||||
 Db 1875 CGGTACCTATCAGGTAGCGCTGAGCAATTG 1845

RESULT 4
 AX275256/c
 LOCUS AX275256 1725 bp DNA linear PAT 29-OCT-2001
 DEFINITION Sequence 5 from Patent WO0171014.
 ACCESSION AX275256
 VERSION AX275256.1 GI:16547676
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (sites)
 AUTHORS Mor, T., Soreq, H., Arntzen, C. and Mason, H.
 TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
 JOURNAL Patent: WO 0171014-A 5 27-SEP-2001;
 BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US); Mor,
 Tsafirir (US); Soreq, Hermona (IL); Arntzen, Charles (US); Mason,
 Hugh S. (US)

FEATURES
 source
 1. 1725
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="synthetic human acetylcholinesterase gene optimized for expression in plants"
 BASE COUNT 273 a 553 c 571 g 328 t
 ORIGIN

Query Match 65.8%; Score 20.4; DB 6; Length 1725;
 Best Local Similarity 95.5%; Pred. No. 1.13e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 10 tcaggtacgcgtgagcaatttg 31
 |||||
 Db 1725 TCAGGTGCGCTGAGCAATTG 1704

RESULT 5
 AC073864/c
 LOCUS AC073864 136257 bp DNA linear PRI 30-MAR-2001
 DEFINITION Homo sapiens 12q BAC RP11-227B21 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
 ACCESSION AC073864
 VERSION AC073864.28 GI:13489126
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 136257)
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Benton, J., Blum, A., Amarantunge, H.C., Are, J.R., Banks, T., Barbara, J., Bowie, S., Brueckner, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burdett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H., Douthwaite, K.J., Draper, H.,

Dugan-Rocha, S., Durbin, K.J., Barnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Haves, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, O., King, L., Korvah, J., Kovar, C., Kratovich, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martine, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwundu, G., Oragunye, N., Oviedo, R., Pacer, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoostari, N., Slisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 136257)
Worley, K.C.
Direct Submission
Submitted (30-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 136257)
Worley, K.C.
Direct Submission
Submitted (30-MAR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 30, 2001 this sequence version replaced gi:13443043.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for

a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu/info/genbank.annotation.html>.

QUALSTAT-REPORT-----

```
----- Summary Statistics -----
Contig length: 136257
Phrap values in estimate: 135295
Average error rate (BCM-Phrap estimate): 0.00010428
Fraction of Phrap values less than 40 : 0.010325
Number of consensus changing edits: 20
Number of N's in consensus : 0
```

```
----- Consensus changing edits -----
Position Original-Context Edited-Context
2846 tgataata(n)caacaataa tgataata(a)caacaataa
3777 agataataat(n)atnttctt agataataat(t)atnttctt
3780 taattatnat(n)ttgtcttct taattatnat(t)ttgtcttct
17066 caaaaaaaca(n)acagatacat caaaaaaaca(t)acagatacat
17341 tgacagtggg(n)acagtcattt tgacagtggg(t)acagtcattt
18186 atcaaaacttt(n)atagaacac atcaaaacttt(a)atagaacac
85229 gtgtcatgt(n)agggtgtgtg gtgtcatgt(g)agggtgtgtg
85348 tgtgtgtga(n)gnngngtctg tgtgtgtga(t)gnngngtctg
85350 tgtgtgtang(n)gnngngtctg tgtgtgtang(g)gnngngtctg
85351 gtgtgtang(n)gnngngtctg gtgtgtang(a)gnngngtctg
85353 gtgtangng(n)ngtctgaagg gtgtangng(g)ngtctgaagg
85355 gtangngng(n)gtctgaggn gtangngng(t)gtctgaggn
85364 gngctcagg(n)ngtngtggag gngctcagg(g)ngtngtggag
85365 ngctcaggn(n)gtngtggaga ngctcaggn(t)gtngtggaga
85369 tgagngngt(n)gtgagaagat tgagngngt(g)gtgagaagat
93974 ttttttttg(g)ggagacaagg ttttttttg(t)ggagacaagg
107496 aaaaaaaaa(n)caaaaacctt aaaaaaaaa(a)caaaaacctt
122522 tctctgtatt(n)tatatacaa tctctgtatt(t)tatatacaa
122608 tggcctaggg(n)ctgttttccc tggcctaggg(t)ctgttttccc
134275 acatgcattgt(n)ccaatcacat acatgcattgt(t)ccaatcacat
```

----- Distribution of Quality < 40 Bases -----

1000	5	10	15	20	25	30	35	40
900								
800								
700								
600								
500								
400								
300								
200								
100								
0	*	*	*	*	*	*	*	*
	phrap Value Range							

```
Version: 1.01 gxf.
Location/Qualifiers
1. .136257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12q"
/clone="RP11-227B21"
repeat_region complement(256..742)
```

```

repeat_region      /rpt_family="L2"
repeat_region      /rpt_family="(CAT)n"
repeat_region      complement(2124..2507)
repeat_region      /rpt_family="MLT1J2"
repeat_region      2654..2805
repeat_region      /rpt_family="MIR"
repeat_region      2826..2862
repeat_region      /rpt_family="(TAA)n"
repeat_region      2978..3059
repeat_region      /rpt_family="MIR"
repeat_region      complement(3057..3372)
repeat_region      /rpt_family="L2"
repeat_region      complement(3435..3731)
repeat_region      /rpt_family="AluX"
repeat_region      3734..3941
repeat_region      /rpt_family="L2"
repeat_region      complement(4185..4267)
repeat_region      /rpt_family="L2"
repeat_region      complement(4367..4527)
repeat_region      /rpt_family="L2"
repeat_region      4528..4830
repeat_region      /rpt_family="AluJb"
repeat_region      complement(4831..4988)
repeat_region      /rpt_family="L2"
repeat_region      complement(5401..5468)
repeat_region      /rpt_family="MIR"
repeat_region      5430..5498
repeat_region      /rpt_family="MIR"
repeat_region      6294..6348
repeat_region      /rpt_family="L2"
repeat_region      complement(6582..6772)
repeat_region      /rpt_family="L2"

Query Match      65.8%; Score 20.4; DB 9; Length 136257;
Best Local Similarity 80.0%; Pred. NO. 80;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ggctaccatcaggtgagcgtgagcaattg 31
||||||| ||||| ||||| |||||
Db 117693 GGTACCTATCTGTAGAGCTGATGATTG 117664

```

```

RESULT 6
AC026678
LOCUS
DEFINITION
AC026678 Homo sapiens chromosome 12 clone RP11-220C5, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION
VERSION AC026678.12 GI:9438510
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Muzny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,
Bodot,B., Bock,J., Bowles,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Foreman-Tansey,J., Frantz,P., Ganes,R., Gorrell,J.H., Gorrell,L.L.,
Guevara-W., Harris,K., Hernandez,J., Hodgson,A., Hoques,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozano,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,N., Nguyen,S.,
Oswald,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucaang,R.,
Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wabnah,M.,
Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,

```

```

Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 170075)
Worley,K.C.
Direct Submission
Submitted (23-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 25, 2000 this sequence version replaced gi:8671821.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HASK
Center clone name: RP11-220C5
----- Summary Statistics
Sequencing vector: M13: L08821
Chemistry: Dye-terminator Big Dye: 5% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 155096 bases at least Q40
Consensus quality: 161754 bases at least Q30
Consensus quality: 164954 bases at least Q20
Estimated insert size: 166552: sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 3.4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 16 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 27634: contig of 27634 bp in length
* 27635 27734: gap of unknown length
* 27735 53075: contig of 25341 bp in length
* 53076 53175: gap of unknown length
* 53176 75203: contig of 22028 bp in length
* 75204 75303: gap of unknown length
* 75304 91881: contig of 16578 bp in length
* 91882 104306: contig of 12325 bp in length
* 104307 104406: gap of unknown length
* 104407 115044: contig of 10638 bp in length
* 115045 115144: gap of unknown length
* 115145 125442: contig of 10298 bp in length
* 125443 134447: contig of 8905 bp in length
* 134448 134547: gap of unknown length
* 134548 141833: contig of 7286 bp in length
* 141834 149333: gap of unknown length
* 149334 149757: contig of 7824 bp in length
* 149758 149857: gap of unknown length
* 149858 154346: contig of 4489 bp in length
* 154347 154446: gap of unknown length
* 154447 158653: contig of 4207 bp in length
* 158654 158753: gap of unknown length
* 158754 163349: contig of 4596 bp in length
* 163350 163449: gap of unknown length
* 163450 166082: contig of 2633 bp in length
* 166083 166182: gap of unknown length
* 166183 168433: contig of 2251 bp in length
* 168434 168533: gap of unknown length
* 168534 170075: contig of 1542 bp in length.
Location/Qualifiers

```

FEATURES

```

source
1. 170075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-220C5"

BASE COUNT 45458 a 38815 c 38968 g 45324 t 1510 others
ORIGIN

Query Match 65.8%; Score 20.4; DB 2; Length 170075;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 gtaactatcagtagcgtgagcaatttg 31
||||| ||||| ||||| ||||| |||||
Db 67708 GGTACTATCTCGTAGAGCTGATAGATTG 67737

RESULT 7
AC024508/c 172241 bp DNA linear HTG 10-SEP-2000
LOCUS Homo sapiens chromosome 1 clone RP11-156A20 map 1, WORKING DRAFT
DEFINITION AC024508
ACCESSION AC024508
VERSION AC024508.4 GI:10047852
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 172241)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 1, clone RP11-156A20
Unpublished
2 (bases 1 to 172241)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bada,F., Boguslavskiy,I.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Hosland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczyk,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivari,T.M.,
Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.

Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 10, 2000 this sequence version replaced gi:8099049.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
-----
Project Information
Center project name: L7204
Center clone name: 156_A_20
-----
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads

```

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 15372 bases at least Q40
Consensus quality: 163469 bases at least Q30
Consensus quality: 166829 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 169541; sum-of-contigs
Quality coverage: 4.3 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 779: contig of 779 bp in length
780 879: gap of 100 bp
880 2388: contig of 1509 bp in length
2389 2488: gap of 100 bp
2489 4256: contig of 1768 bp in length
4257 4356: gap of 100 bp
4357 5784: contig of 1428 bp in length
5785 5884: gap of 100 bp
5885 7892: contig of 2008 bp in length
7893 7992: gap of 100 bp
7993 9655: contig of 1663 bp in length
9656 9755: gap of 100 bp
9756 12205: contig of 2450 bp in length
12206 12305: gap of 100 bp
12306 14310: contig of 2005 bp in length
14311 14410: gap of 100 bp
14411 16707: contig of 2297 bp in length
16708 16807: gap of 100 bp
16808 19279: contig of 2472 bp in length
19280 19379: gap of 100 bp
19380 21937: contig of 2558 bp in length
21938 22037: gap of 100 bp
22038 48726: contig of 26689 bp in length
48727 48826: gap of 100 bp
48827 51847: contig of 3021 bp in length
51848 51947: gap of 100 bp
51948 54626: contig of 2679 bp in length
54627 54726: gap of 100 bp
54727 58944: contig of 4218 bp in length
58945 59044: gap of 100 bp
59045 63316: contig of 4272 bp in length
63317 63416: gap of 100 bp
63417 68257: contig of 4841 bp in length
68258 68357: gap of 100 bp
68358 73280: contig of 4923 bp in length
73281 73380: gap of 100 bp
73381 78209: contig of 4829 bp in length
78210 78309: gap of 100 bp
78310 84542: contig of 6233 bp in length
84543 84642: gap of 100 bp
84643 89899: contig of 5257 bp in length
89900 89999: gap of 100 bp
90000 96909: contig of 6910 bp in length
96910 97009: gap of 100 bp
97010 105708: contig of 8699 bp in length
105709 105808: gap of 100 bp
105809 116467: contig of 10659 bp in length
116468 116567: gap of 100 bp
116568 128407: contig of 11840 bp in length
128408 128507: gap of 100 bp
128508 144528: contig of 16021 bp in length
144529 144628: gap of 100 bp
144629 169246: contig of 24618 bp in length
169247 169346: gap of 100 bp
169347 172241: contig of 2895 bp in length.

SOURCE INFORMATION:

The RPO1-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanesi, J. G. and de Jong, P. J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-12616. Actual start of this clone is at base position 1 of RP11-431K24; actual end is at base position 178820 of RP11-431K24.

The fidelity of the dinucleotide (GA) repeat from base position 123460 to 123487 cannot be guaranteed. The region at base position 123460 to 1234624 is covered by a single subclone. The fidelity of sequence at base position 124152 to 124176 cannot be guaranteed. The regions from 123927 to 123930 and 145393 to 145403 are single-stranded and sequenced with only one chemistry.

FEATURES

source	1. .178820 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1" /map="1" /clone="RP11-431K24" /clone_lib="RPC1-11" 5. .71 /rpt_family="L1" 87. .364 /rpt_family="Alu" 1272. .1553 /rpt_family="Alu" 2779. .3084 /rpt_family="Alu" 3739. .4048 /rpt_family="Alu" 4229. .4444 /rpt_family="MIR" 4540. .4846 /rpt_family="Alu" 4886. .5014 /rpt_family="ERV1" 5113. .5372 /rpt_family="ERV1" 6568. .7005 /note="similar to EST AI022484 (NID:g33237725) ow50c05.s1" 6720. .7031 /rpt_family="Alu" 7097. .7212 /rpt_family="L2" 7112. .7473 /note="similar to EST AI141387 (NID:g3648844) oy41f08.s1" 7158. .7444 /note="similar to EST AA843283 (NID:g2929801) aj16d01.s1" 7546. .7650 /rpt_family="L1" 7651. .8191 /rpt_family="ERV1" 8192. .8374 /rpt_family="L1" 8353. .8404 /rpt_family="L1" 8407. .8882 /rpt_family="ERV1" 8887. .9198 /rpt_family="Alu" 9202. .9464 /rpt_family="ERV1" 9471. .9595
--------	--

repeat_region	/rpt_family="L1"	9610, .9711
repeat_region	/rpt_family="L1"	9720, .10077
repeat_region	/rpt_family="L1"	10078, .10375
repeat_region	/rpt_family="Alu"	10376, .10543
repeat_region	/rpt_family="L1"	10997, .11038
repeat_region	/rpt_family="MaLR"	11039, .11342
repeat_region	/rpt_family="Alu"	11343, .11651
repeat_region	/rpt_family="MaLR"	11775, .12137
repeat_region	/rpt_family="MaLR"	12709, .13030
repeat_region	/rpt_family="Alu"	13183, .13482
repeat_region	/rpt_family="Alu"	13660, .13957
repeat_region	/rpt_family="Alu"	14067, .14300
repeat_region	/rpt_family="MaLR"	14301, .14602
repeat_region	/rpt_family="Alu"	14603, .14782
repeat_region	/rpt_family="MaLR"	14969, .15266
repeat_region	/rpt_family="Alu"	15341, .15642
repeat_region	/rpt_family="Alu"	15657, .15960
repeat_region	/rpt_family="Alu"	16373, .16661
repeat_region	/rpt_family="Alu"	16709, .16850
repeat_region	/rpt_family="L1"	17101, .17178
repeat_region	/rpt_family="ERV1"	17179, .17499
repeat_region	/rpt_family="Alu"	17500, .17908
repeat_region	/rpt_family="ERV1"	17921, .18009
repeat_region	/rpt_family="Alu"	18064, .18195
repeat_region	/rpt_family="Alu"	19470, .19669
repeat_region	/rpt_family="L1"	19765, .20070
repeat_region	/rpt_family="ERV1"	20369, .20670
repeat_region	/rpt_family="Alu"	21550, .21850
repeat_region	/rpt_family="Alu"	22158, .22462
repeat_region	/rpt_family="Alu"	23533, .23834
repeat_region	/rpt_family="Alu"	23886, .24012
repeat_region	/rpt_family="Alu"	24308, .24637

Query Match 65.8%; Score 20.4; DB 9; Length 178820;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0;

Qy 2 ggtaacctatcaggtagcgctgagcaatttg 31
 ||| ||||| ||| ||| ||| ||| |
Db 106290 GGTTCCTATCAGCAAGCGATGACCAATTAG 106261

```

RESULT 9
AL358876/c 181230 bp DNA llinear HTG 28-NOV-2001
LOCUS Homo sapiens chromosome 1 clone RP11-431K24, *** SEQUENCING IN
DEFINITION PROGRESS ***, 6 unordered pieces.
ACCESSION AL358876
VERSION AL358876
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hammond,S.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Nov 29, 2001 this sequence version replaced gi:16944018.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA31K24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 180142 bases at least Q40
Consensus quality: 180483 bases at least Q30
Consensus quality: 180640 bases at least Q20
Insert size: 180730; sum-of-contigs
Insert size: 190919; 1.1% error; agarose-fp
Quality coverage: 8.84x in Q20 bases; sum-of-contigs Quality
coverage: 8.42x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 31912: contig of 31912 bp in length
* 31913 32012: gap of 100 bp
* 32013 114102: contig of 82090 bp in length
* 114103 114202: gap of 100 bp
* 114203 128075: contig of 13873 bp in length
* 128076 128175: gap of 100 bp
* 128176 146992: contig of 18817 bp in length
* 146993 147092: gap of 100 bp
* 147093 161684: contig of 14592 bp in length
* 161685 161784: gap of 100 bp
* 161785 181230: contig of 19446 bp in length.
*
* Location/Qualifiers
* 1. .181230
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="1"
* /clone="RP11-431K24"
* /clone_lib="RPCI-11.2"
* 1. .31912
* /note="assembly_fragment:02646"
* fragment_chain:1"
* 32013..114102
* /note="assembly_fragment:02199"
* fragment_chain:1"
* 114203..128075
* /note="assembly_fragment:01442"

```

```

fragment_chain:1"
128176..146992
/note="assembly_fragment:00582"
fragment_chain:1"
147093..161684
/note="assembly_fragment:00644"
fragment_chain:1"
161785..181230
/note="assembly_fragment:00898"
clone_end:SP6
vector_side:right"
BASE COUNT 47481 a 41685 c 43504 g 48060 t 500 others
ORIGIN
Query Match 65.8%; Score 20.4; DB 2; Length 181230;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 ggtacctatcagtagcgctgagcaatttg 31
||||| ||||||| ||||| ||| ||||| |
Db 110227 GGTTCCTATCAGCAGCGATGACCAATTAG 110198

RESULT 10
AC023183/c 183754 bp DNA llinear HTG 25-MAY-2000
LOCUS Homo sapiens clone RP11-677L3, WORKING DRAFT SEQUENCE, 36 unordered
DEFINITION pieces.
ACCESSION AC023183
VERSION AC023183.3 GI:8072560
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 183754)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-677L3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 183754)
AUTHORS Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chospel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo.A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On May 25, 2000 this sequence version replaced gi:7139681.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4323
Center clone name: 677_L_3

```

```
----- Summary Statistics
Sequencing vector: M13; W7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162428 bases at least Q40
Consensus quality: 171987 bases at least Q30
Consensus quality: 176412 bases at least Q20
Insert size: 192000; agarose-fp
Insert size: 180254; sum-of-contigs
Quality coverage: 3.5 in Q20 bases; agarose-fp
Quality coverage: 3.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1085: contig of 1085 bp in length
* 1086 1185: gap of 100 bp
* 1186 2242: contig of 1057 bp in length
* 2243 2342: gap of 100 bp
* 2343 3408: contig of 1066 bp in length
* 3409 3508: gap of 100 bp
* 3509 3956: contig of 448 bp in length
* 3957 4056: gap of 100 bp
* 4057 5612: contig of 1556 bp in length
* 5613 5712: gap of 100 bp
* 5713 7899: contig of 2187 bp in length
* 7900 7999: gap of 100 bp
* 8000 9476: contig of 1477 bp in length
* 9477 9576: gap of 100 bp
* 9577 11577: contig of 2001 bp in length
* 11578 11677: gap of 100 bp
* 11678 13800: contig of 2123 bp in length
* 13801 13900: gap of 100 bp
* 13901 15510: contig of 1610 bp in length
* 15511 15610: gap of 100 bp
* 15611 17280: contig of 1670 bp in length
* 17281 17380: gap of 100 bp
* 17381 19596: contig of 2216 bp in length
* 19597 19696: gap of 100 bp
* 19697 21776: contig of 2080 bp in length
* 21777 21876: gap of 100 bp
* 21877 24032: contig of 2156 bp in length
* 24033 24132: gap of 100 bp
* 24133 26500: contig of 2518 bp in length
* 26501 26750: gap of 100 bp
* 26751 30914: contig of 4164 bp in length
* 30915 31014: gap of 100 bp
* 31015 34039: contig of 3025 bp in length
* 34040 34139: gap of 100 bp
* 34140 37599: contig of 3460 bp in length
* 37600 37699: gap of 100 bp
* 37700 41746: contig of 4047 bp in length
* 41747 41846: gap of 100 bp
* 41847 46449: contig of 4603 bp in length
* 46450 46549: gap of 100 bp
* 46550 51055: contig of 4506 bp in length
* 51056 51155: gap of 100 bp
* 51156 55588: contig of 4433 bp in length
* 55589 55688: gap of 100 bp
* 55689 61100: contig of 5412 bp in length
* 61101 61200: gap of 100 bp
* 61201 65672: contig of 4472 bp in length
* 65673 65772: gap of 100 bp
* 65773 72053: contig of 6281 bp in length
* 72054 72153: gap of 100 bp
* 72154 76428: contig of 4275 bp in length
* 76429 76528: gap of 100 bp
* 76529 84178: contig of 7650 bp in length
*
* 84179 84278: gap of 100 bp
* 84279 92118: contig of 7840 bp in length
* 92119 92218: gap of 100 bp
* 92219 99084: contig of 6866 bp in length
* 99085 99184: gap of 100 bp
* 99185 107991: contig of 8807 bp in length
* 107992 108091: gap of 100 bp
* 108092 116709: contig of 8618 bp in length
* 116710 116809: gap of 100 bp
* 116810 128617: contig of 11808 bp in length
* 128618 128717: gap of 100 bp
* 128718 142241: contig of 13524 bp in length
* 142242 142341: gap of 100 bp
* 142342 154807: contig of 12466 bp in length
* 154808 154907: gap of 100 bp
* 154908 167742: contig of 12835 bp in length
* 167743 167842: gap of 100 bp
* 167843 183754: contig of 15912 bp in length.
*
FEATURES
  source
    1..183754
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="RP11-677L3"
      /clone_lib="RPC1-11 Human Male BAC"
    1..1085
      /note="assembly_fragment"
    1186..2242
      /note="assembly_fragment"
    2343..3408
      /note="assembly_fragment"
    3509..3956
      /note="assembly_fragment"
      clone_end:T7
      vector_side:right
    4057..5612
      /note="assembly_fragment"
    5713..7899
      /note="assembly_fragment"
    8000..9476
      /note="assembly_fragment"
    9577..11577
      /note="assembly_fragment"
    11678..13800
      /note="assembly_fragment"
    13901..15510
      /note="assembly_fragment"
    15611..17280
      /note="assembly_fragment"
    17381..19596
      /note="assembly_fragment"
    19697..21776
      /note="assembly_fragment"
    21877..24032
      /note="assembly_fragment"
    24133..26500
      /note="assembly_fragment"
    26751..30914
      /note="assembly_fragment"
    31015..34039
      /note="assembly_fragment"
    34140..37599
      /note="assembly_fragment"
    37700..41746
      /note="assembly_fragment"
    41847..46449
      /note="assembly_fragment"
    46550..51055
      /note="assembly_fragment"
    51156..55588
      /note="assembly_fragment"
    55689..61100
      /note="assembly_fragment"
    61201..65672
      /note="assembly_fragment"
```

```

/misc_feature /note="assembly_fragment"
65773..72053
/misc_feature /note="assembly_fragment"
72154..76428
/note="assembly_fragment
clone_end:SP6
vector_side:right"

Query Match 65.8%; Score 20.4; DB 2; Length 183754;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ggtacattacaggtagcgtgagcaatttg 31
||||| ||||| ||||| |||||
DB 136461 GGTACCTATCTCGTAGAGCTGATGATTG 136432

RESULT 11
AC024935/C
DEFINITION Homo sapiens chromosome 12 clone RP11-471G7, WORKING DRAFT
AC024935
AC024935 260430 bp DNA linear HTG 15-OCT-2001
VERSION AC024935.29 GI:16117894
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 260430)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Isman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bivage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P., Buhsy,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyte,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S.,
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
Slisdon,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczek,R., Woden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 260430)
Worley,K.C.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (03-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 14, 2001 this sequence version replaced gi:11995488.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAJN
Center clone name: RP11-471G7
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye: 99% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 260230 bases at least Q40
Consensus quality: 286983 bases at least Q30
Consensus quality: 302060 bases at least Q20
Estimated insert size: 285925; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 3.8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
consists of 28 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 62329: contig of 62329 bp in length
* 62429: gap of unknown length
* 62430: contig of 43264 bp in length
* 105693: gap of unknown length
* 105794: contig of 33355 bp in length
* 139149: gap of unknown length
* 139248: gap of unknown length
* 165659: contig of 26411 bp in length
* 165759: gap of unknown length
* 174283: contig of 8524 bp in length
* 174284: gap of unknown length
* 174384: contig of 6831 bp in length
* 181215: gap of unknown length
* 181315: contig of 7108 bp in length
* 188422: gap of unknown length
* 188522: contig of 5163 bp in length
* 193685: gap of unknown length
* 193785: gap of unknown length
* 193786: contig of 5873 bp in length
* 199659: gap of unknown length
* 199759: gap of unknown length
* 204510: contig of 4752 bp in length
* 204610: gap of unknown length
* 208759: contig of 4149 bp in length
* 208859: gap of unknown length
* 208860: contig of 4651 bp in length
* 213510: gap of unknown length
* 213610: contig of 3559 bp in length
* 217169: gap of unknown length
* 217289: contig of 3166 bp in length
* 220435: gap of unknown length
* 220535: contig of 4135 bp in length
* 224670: gap of unknown length
* 224771: contig of 3045 bp in length
* 227815: gap of unknown length
* 227915: contig of 3080 bp in length
* 230995: gap of unknown length
* 231095: contig of 2079 bp in length
* 231174: gap of unknown length
* 233175: contig of 3365 bp in length
* 236639: contig of 3365 bp in length

```



```

* 236640 236739: gap of unknown length
* 236740 240079: contig of 3340 bp in length
* 240080 240179: gap of unknown length
* 240180 242733: contig of 2554 bp in length
* 242734 242833: gap of unknown length
* 242834 245854: contig of 3021 bp in length
* 245855 245954: gap of unknown length
* 245955 248397: contig of 2443 bp in length
* 248398 248497: gap of unknown length
* 248498 251331: contig of 2834 bp in length
* 251332 251431: gap of unknown length
* 251432 253554: contig of 2123 bp in length
* 253555 253654: gap of unknown length
* 253655 253726: contig of 2072 bp in length
* 253727 255826: gap of unknown length
* 255827 257863: contig of 2037 bp in length
* 257864 257963: gap of unknown length
* 257964 260430: contig of 2467 bp in length.
FEATURES
    Location/Qualifiers
        1..260430
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="12"
            /clone="RP11-471G7"
BASE COUNT      78232 a 52577 c 53495 g 73388 t 2738 others
ORIGIN

Query Match      65.8%; Score 20.4; DB 2; Length 260430;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtagcgtgagcaatttg 31
Db 187395 GGTACCTATCTCGTAGAGCTGATAGATTG 187366

RESULT 12
DSP276564
LOCUS      Desulfotomaculum sp. DEM-KMe99-1 1019 bp rRNA linear BCT 19-NOV-2000
DEFINITION Desulfotomaculum sp. DEM-KMe99-1 16S rRNA gene.
ACCESSION AJ276564
VERSION AJ276564.1 GI:11230398
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.
SOURCE Desulfotomaculum sp. DEM-KMe99-1.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Heliobacterium
group; Desulfotomaculum; environmental samples.
REFERENCE 1 (bases 1 to 1019)
AUTHORS Stubner,S. and Meuser,K.
TITLE Detection of Desulfotomaculum in an Italian rice paddy soil by 16S
ribosomal nucleic acid analyses
JOURNAL FEMS Microbiol. Ecol. 34 (1), 73-80 (2000)
PUBMED 11053738
REFERENCE 2 (bases 1 to 1019)
AUTHORS Stubner,S.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Stubner S., Biogeochemie,
Max-Planck-Institut f. terrestrische Mikrobiologie,
Karl-von-Frisch-Strasse, 35043 Marburg, GERMANY
FEATURES
    Location/Qualifiers
        1..1019
            /organism="Desulfotomaculum sp. DEM-KMe99-1"
            /db_xref="taxon:118333"
            /clone="DEM-KMe99-1"
            /note="isolated from rice paddy soil"
rRNA      1..1019
            /gene="16S rRNA"
            /product="16S ribosomal RNA"
            1..1019
            /gene="16S rRNA"
BASE COUNT      257 a 228 c 339 g 194 t 1 others
ORIGIN

```

```

Query Match      64.5%; Score 20; DB 1; Length 1019;
Best Local Similarity 82.1%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 gtacctatcaggtagcgtgagcaattt 30
Db 978 GTTGCTAACAGGTAGCGCTGAGCACTCT 1005

RESULT 13
AB015893
LOCUS      Treponema socranskii subsp. buccale gene for 16S rRNA, partial
DEFINITION sequence, isolate:T10.
ACCESSION AB015893
VERSION AB015893.1 GI:4150946
KEYWORDS 16S ribosomal RNA.
SOURCE Treponema socranskii subsp. buccale (sub_species:buccale,
isolate:T10) DNA.
ORGANISM Treponema socranskii subsp. buccale
REFERENCE 1 (sites)
AUTHORS Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and
Nakase,T.
TITLE Phylogenetic analysis of saccharolytic oral treponemes isolated
from human subgingival plaque
JOURNAL Microbiol. Immunol. 43 (7), 711-716 (1999)
MEDLINE 99456392
REFERENCE 2 (bases 1 to 1409)
AUTHORS Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and
Nakase,T.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1998) to the DDBJ/EMBL/GenBank databases. Mitsuo
Sakamoto, The Institute of Physical and Chemical Research (RIKEN),
Japan Collection of Microorganisms; Hirotsawa 2-1, Wako, Saitama
351-0198, Japan (E-mail:sakamoto@elmus.riken.go.jp,
Tel:81-48-462-1111(ex.5136), Fax:81-48-462-4619,
On Jan 12, 1999 this sequence version replaced gi:4115460.
FEATURES
    Location/Qualifiers
        1..1409
            /organism="Treponema socranskii subsp. buccale"
            /isolate="T10"
            /sub_species="buccale"
            /db_xref="taxon:69713"
            <1..>1409
            /product="16S rRNA"
BASE COUNT      353 a 339 c 483 g 234 t
ORIGIN

Query Match      64.5%; Score 20; DB 1; Length 1409;
Best Local Similarity 82.1%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 cggttacctatcaggtagcgtgagcaat 28
Db 1064 CGGTACTTAACAGGTAGCGCTGAGCACT 1091

RESULT 14
AB015888
LOCUS      Treponema socranskii subsp. socranskii gene for 16S rRNA, partial
DEFINITION sequence, isolate:1166.
ACCESSION AB015888
VERSION AB015888.1 GI:4150941
KEYWORDS 16S ribosomal RNA.
SOURCE Treponema socranskii subsp. socranskii (sub_species:socranskii,
isolate:1166) DNA.
ORGANISM Treponema socranskii subsp. socranskii
Bacteria; Spirochaetales; Spirochaetaceae; Treponema.

```

```

REFERENCE
AUTHORS      1 (sites)
              Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and
              Nakase,T.
TITLE        Phylogenetic analysis of saccharolytic oral treponemes isolated
              from human subgingival plaque
JOURNAL      Microbiol. Immunol. 43 (7), 711-716 (1999)
MEDLINE      99456392
REFERENCE    2 (bases 1 to 1410)
AUTHORS      Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and
              Nakase,T.
TITLE        Direct Submission
JOURNAL      Submitted (30-JUN-1998) to the DDBJ/EMBL/GenBank databases. Mitsuo
              Sakamoto, The Institute of Physical and Chemical Research (RIKEN),
              Japan Collection of Microorganisms; Hirogawa 2-1, Wako, Saitama
              351-0198, Japan (E-mail:sakamoto@ulmus.riken.go.jp,
              Tel:81-48-462-1111(ex.5136), Fax:81-48-462-4619)
COMMENT      On Jan 12, 1999 this sequence version replaced gi:4115455.
FEATURES     Location/Qualifiers
              source
                1..1410
                /organism="Treponema socranskii subsp. socranskii"
                /isolate="1166"
                /sub_species="socranskii"
                /db_xref="taxon:69715"
                <1..>1410
                /product="16S rRNA"
BASE COUNT   354 a 335 c 485 g 236 t
ORIGIN
rRNA
Query Match      64.5%; Score 20; DB 1; Length 1410;
Best Local Similarity 82.1%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cggctacatcaggtagcgctgagcaat 28
    |||| ||| ||||| ||||| |||
Db 1066 CGGTTACTACAGGTAGCGCTGAGGACT 1093

RESULT 15
AB015889
LOCUS      AB015889                1414 bp    DNA    linear    BCT 31-JUL-1999
DEFINITION Treponema socranskii subsp. socranskii gene for 16S rRNA, partial
              sequence, isolate:128B.
ACCESSION  AB015889
VERSION    AB015889.1 GI:4150942
KEYWORDS   16S ribosomal RNA.
SOURCE     Treponema socranskii subsp. socranskii (sub_species:socranskii,
              isolate:128B) DNA.
ORGANISM   Treponema socranskii subsp. socranskii
REFERENCE  Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
AUTHORS    Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and
              Nakase,T.
TITLE      Phylogenetic analysis of saccharolytic oral treponemes isolated
              from human subgingival plaque
JOURNAL    Microbiol. Immunol. 43 (7), 711-716 (1999)
MEDLINE    99456392
REFERENCE  2 (bases 1 to 1414)
AUTHORS    Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and
              Nakase,T.
TITLE      Direct Submission
JOURNAL    Submitted (30-JUN-1998) to the DDBJ/EMBL/GenBank databases. Mitsuo
              Sakamoto, The Institute of Physical and Chemical Research (RIKEN),
              Japan Collection of Microorganisms; Hirogawa 2-1, Wako, Saitama
              351-0198, Japan (E-mail:sakamoto@ulmus.riken.go.jp,
              Tel:81-48-462-1111(ex.5136), Fax:81-48-462-4619)
COMMENT    On Jan 12, 1999 this sequence version replaced gi:4115456.
FEATURES   Location/Qualifiers
              source
                1..1414
                /organism="Treponema socranskii subsp. socranskii"
                /isolate="128B"
                /sub_species="socranskii"
                /db_xref="taxon:69715"

```

```

rRNA
BASE COUNT   354 a 337 c 484 g 239 t
ORIGIN
rRNA
Query Match      64.5%; Score 20; DB 1; Length 1414;
Best Local Similarity 82.1%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cggctacatcaggtagcgctgagcaat 28
    |||| ||| ||||| ||||| |||
Db 1069 CGGTTACTACAGGTAGCGCTGAGGACT 1096

Search completed: August 31, 2002, 18:04:30
Job time: 15723 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:17:15 ; Search time 824.08 Seconds
(without alignments)
64.586 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtagctatcaggtagcgtgagcaatttg 31

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	31	24	Human acetylcholin
c 2	31	100.0	5767	24	Plasmid pTM034. S
c 3	31	100.0	14445	24	Plasmid pTM036. S
c 4	20.4	65.8	1725	24	Synthetic human ac
5	20	64.5	256	21	Human secreted pro
6	20	64.5	449	22	Human breast cell
7	20	64.5	449	22	Human foetal liver
8	20	64.5	449	22	Human foetal liver
9	20	64.5	449	22	Probe #3661 for 9e

10	20	64.5	449	22	ABA25248	Probe #3714 for 9e
11	20	64.5	449	22	AAK29178	Human bone marrow
12	20	64.5	449	22	AAK29237	Human bone marrow
13	20	64.5	449	22	AAI13769	Probe #3702 for 9e
14	20	64.5	449	22	AAI13830	Probe #3763 for 9e
15	20	64.5	449	22	AAI35130	Probe #3816 used t
16	20	64.5	449	22	AAI03648	Human polynucleoti
c 17	20	1897	5476	22	AAI59644	Human polynucleoti
18	19	61.3	5476	22	ABAI8966	Human nervous syst
19	19	61.3	5476	22	ABAI8967	Human nervous syst
20	19	61.3	5476	22	AAI36446	Human musculoskele
21	19	61.3	5476	22	AAI36447	Human musculoskele
22	19	61.3	5476	22	AAI32794	Human genomic DNA
23	19	61.3	5476	22	AAI32795	Human genomic DNA
24	19	61.3	5476	22	AAI326816	Human genomic DNA
25	19	61.3	5476	22	AAI326817	Human genomic DNA
c 26	18.8	60.6	5471	24	AB199407	Mouse ischaemic co
27	18.8	60.6	16075	20	AAV99811	Gun gene cluster.
28	18.8	60.6	16079	8	AAV07553	Sequence of segmen
29	18.2	58.7	47	21	AAZ87020	RBP-7 biallelic ma
30	18.2	58.7	1366	21	AAA15481	16S rRNA sequence
c 31	18.2	58.7	4875	23	ABL16962	Drosophila melanog
32	18	58.1	1104	23	ABL20007	Drosophila melanog
33	18	58.1	3278	23	ABL20006	Drosophila melanog
c 34	17.8	57.4	879	21	AAI45814	Arabidopsis thalia
c 35	17.8	57.4	1035	21	AAI41930	Arabidopsis thalia
c 36	17.8	57.4	3573	23	ABL02260	Drosophila melanog
c 37	17.8	57.4	3948	23	ABL03664	Drosophila melanog
c 38	17.8	57.4	6065	22	AAI32886	Human genomic DNA
c 39	17.8	57.4	12113	24	AAI17492	Human acetylcholin
c 40	17.6	56.8	669	21	AAI12483	Aspergillus oryzae
c 41	17.6	56.8	7461	22	AAI30639	DNA encoding novel
c 42	17.6	56.8	7461	22	AAI28701	Genomic sequence #
c 43	17.6	56.8	32174	22	ABAI5665	Human nervous syst
c 44	17.6	56.8	32174	22	ABAI9477	Human nervous syst
45	17.6	56.8	32174	22	ABA20359	Human nervous syst

ALIGNMENTS

RESULT 1
AAI17546 standard; DNA; 31 BP.

XX AAI17546;

XX AAI17546;

XX 25-FEB-2002 (first entry)

XX Human acetylcholinesterase, PCR primer AChE-Kpn.

Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
transgenic plant; acetylcholinesterase poisoning; chemical warfare;
muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
organophosphate(OP)-modified AChE; pyridostigmine bromide; ss.

XX Homo sapiens.

OS Synthetic.

XX WO200171014-A2.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08468.

XX 17-MAR-2000; 2000US-190440P.

XX (MORT/) MOR T.

XX (SORE/) SOREQ H.

XX (ARNT/) ARNTZEN C.

XX (MASO/) MASON H.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI: 2002-055120/07.
XX
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
XX Claim 11; Page 4; 42pp; English.
XX
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents PCR primer AChE-kpn used to amplify
CC human AChE as described in the method of the invention.
XX
XX Sequence 31 BP; 7 A; 7 C; 9 G; 8 T; 0 other;
SQ

Query Match 100.0%; Score 31; DB 24; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cggctacctatcaggtagcgctgagcaatttg 31
|||||
Db 1 cggctacctatcaggtagcgctgagcaatttg 31

RESULT 2
AAS17547/C
ID AAS17547 standard; DNA; 5767 BP.
XX
XX AAS17547;
XX
XX 25-FEB-2002 (first entry)
XX
XX Plasmid pTW034.
XX
XX Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTW034;
KW circular; ds.
XX
XX Synthetic.
OS
XX WO200171014-A2.
XX
XX 27-SEP-2001.
XX
XX 16-MAR-2001; 2001WO-US08468.
XX
XX 17-MAR-2000; 2000US-190440P.
XX
XX (MORT/) MOR T.
PA

(SORE/) SOREQ H.
(ARNT/) ARNTZEN C.
(MASON/) MASON H.
(BOYC-) BOYCE THOMPSON INST PLANT RES INC.
PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI: 2002-055120/07.
XX
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
XX Claim 11; Page 28-31; 42pp; English.
XX
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTW034, the pcpivkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
XX Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;
SQ

Query Match 100.0%; Score 31; DB 24; Length 5767;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cggctacctatcaggtagcgctgagcaatttg 31
|||||
Db 2571 CGGTACCTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
AAS17548/C
ID AAS17548 standard; DNA; 14446 BP.
XX
XX AAS17548;
XX
XX 25-FEB-2002 (first entry)
XX
XX Plasmid pTW036.
XX
XX Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTW036;
KW circular; ds.
XX
XX Synthetic.
OS
XX WO200171014-A2.
XX
XX 27-SEP-2001.
PD

```
XX 16-MAR-2001; 2001WO-US08468.
XX 17-MAR-2000; 2000US-190440P.
XX (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 32-41; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM036, the pGPTvkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;

Query Match 100.0%; Score 31; DB 24; Length 14446;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtagcctcaggtgagcgtgagcaattg 31
|||||
DB 1875 CGGTACCTATCAGGTAGCCCTGAGCAATTG 1845

RESULT 4
AAS17549/c
ID AAS17549 standard; DNA; 1725 BP.
XX
AC AAS17549;
XX
XX 25-FEB-2002 (first entry)
DE Synthetic human acetylcholinesterase gene.
XX
XX Human; acetylcholinesterase; AChE; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.
XX

OS Homo sapiens.
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
XX 16-MAR-2001; 2001WO-US08468.
XX
XX 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX
XX WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 41-42; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents synthetic human AChE gene optimised for
CC expression in plants.
XX
SQ Sequence 1725 BP; 273 A; 553 C; 571 G; 328 T; 0 other;

Query Match 65.8%; Score 20.4; DB 24; Length 1725;
Best Local Similarity 95.5%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 tcaggtagcgtgagcaattg 31
|||||
DB 1725 TCAGGTGGCGCTGAGCAATTG 1704

RESULT 5
AAC29843
ID AAC29843 standard; cDNA; 256 BP.
XX
AC AAC29843;
XX
XX 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 33918.
XX
```



```

AAK29237
ID AAK29237 standard; DNA; 449 BP.
XX
AC AAK29237;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 3794.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PD WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 3702; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 tacctatcaggtagcgtgagcaatttg 31
| ||| ||||| ||||| |||||
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 14
AAI13830
ID AAI13830 standard; DNA; 449 BP.
XX
AC AAI13830;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #3763 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
XX
AAK29237
ID AAK29237 standard; DNA; 449 BP.
XX
AC AAK29237;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 3794.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PD WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 3794; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 tacctatcaggtagcgtgagcaatttg 31
| ||| ||||| ||||| |||||
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 13
AAI13769
ID AAI13769 standard; DNA; 449 BP.
XX
AC AAI13769;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #3702 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX

```

```

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID No 3763; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENP). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match      64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 taccatcaggtagcgctgagcaatttg 31
   | | | | | | | | | | | | | | | |
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 15
AAI35130
ID AAI35130 standard; DNA; 449 BP.
XX
XX AAI35130;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #3816 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 3816; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match      64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 taccatcaggtagcgctgagcaatttg 31
   | | | | | | | | | | | | | | | |
Db 196 ttctcccaggtagcactgagcagtttg 223

Search completed: August 31, 2002, 18:17:17
Job time: 16439 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:24:33 ; Search time 6064.26 Seconds
(without alignments)
68.995 Million cell updates/sec

Title: US-09-810-861b-2
Perfect score: 31
Sequence: 1 cggtaacctatcaggtagcgcgagcaatttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21.4	69.0	419	10 BF082150	BF082150 QVO-BN004
C 2	19.8	63.9	574	12 BH400848	BH400848 AG-ND-143
C 3	19.8	63.9	687	12 BH398899	BH398899 AG-ND-159
C 4	19.8	63.9	820	10 BF786352	BF786352 602112868
5	19.8	63.9	866	12 CNS01Y2C	AL172465 Tetradon
6	19.4	62.6	430	10 BG984091	BG984091 PM3-CN015
7	19.4	62.6	664	9 AF201206	AF201206 AF201206
C 8	19.4	62.6	668	10 BE370901	BE370901 601219331
C 9	19.4	62.6	776	10 BG821572	BG821572 602725163
10	19.2	61.9	646	10 BE534304	BE534304 601235365
C 11	19	61.3	316	9 BB852642	BB852642 BB852642
C 12	19	61.3	361	9 AU005717	AU005717 AU005717
13	19	61.3	379	10 BG228080	BG228080 ux47d04.x
C 14	19	61.3	390	10 BE634195	BE634195 uv70b03.y
15	19	61.3	390	10 BE986391	BE986391 UI-M-CG0p
16	19	61.3	395	10 BE692000	BE692000 uv70b03.x
17	19	61.3	403	9 AA959441	AA959441 vw59g04.s

18	19	61.3	417	12 AQ789846	AQ789846 HS_3091_B
19	19	61.3	429	9 AI850006	AI850006 UI-M-BG0-
C 20	19	61.3	437	9 AA170254	AA170254 ms87f10.f
21	19	61.3	447	9 AI854861	AI854861 UI-M-BH0-
22	19	61.3	459	9 AI850010	AI850010 UI-M-BG0-
C 23	19	61.3	465	9 AU005700	AU005700 AU005700
C 24	19	61.3	488	10 BG277664	BG277664 ux47d04.y
25	19	61.3	504	9 AI646695	AI646695 ub64f12.x
26	19	61.3	517	10 BF455638	BF455638 UI-M-CG0p
27	19	61.3	538	9 AI840147	AI840147 UI-M-A00-
28	19	61.3	538	10 BF461699	BF461699 UI-M-CG0p
C 29	19	61.3	542	9 AU004943	AU004943 AU004943
C 30	19	61.3	543	9 AU006277	AU006277 AU006277
C 31	19	61.3	566	9 AV398885	AV398885 AV398885
C 32	19	61.3	577	9 AV402228	AV402228 AV402228
33	19	61.3	584	10 BE988805	BE988805 UI-M-CG0p
34	19	61.3	584	10 BF464864	BF464864 UI-M-CG0p
35	19	61.3	585	10 BG228490	BG228490 ux44b05.x
36	19	61.3	585	10 BE623027	BE623027 up51b04.x
37	19	61.3	600	10 BG808869	BG808869 2123-34 M
C 38	19	61.3	604	9 AV405785	AV405785 AV405785
39	19	61.3	605	9 BB619072	BB619072 BB619072
40	19	61.3	625	10 BF465972	BF465972 UI-M-CG0p
C 41	19	61.3	638	9 AV402654	AV402654 AV402654
C 42	19	61.3	645	9 AU006433	AU006433 AU006433
C 43	19	61.3	645	9 AV405236	AV405236 AV405236
C 44	19	61.3	646	9 AU006241	AU006241 AU006241
45	19	61.3	646	9 BB641643	BB641643 BB641643

ALIGNMENTS

RESULT	1
BF082150	419 bp mRNA linear EST 18-OCT-2000
LOCUS	QVO-BN0042-150900-389-e07 BN0042 Homo sapiens cDNA, mRNA sequence.
DEFINITION	BF082150
ACCESSION	BF082150.1 GI:10875980
VERSION	EST.
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	1 (bases 1 to 419)
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baig,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPES/LICR Human Cancer Genome project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QVO-BN0042-150900-389-e07&ts=2000-09-15&tl=1) Seq primer: puc 18 forward High quality sequence stop: 419. Location/Qualifiers 1..419 /organism="Homo sapiens" /db_xref="taxon:9606"

Qy	1	cggtagcctatcaagstagcgctgagcaatttg	31
Db	216	CGGAACATATCAGTAGCTGTAAACAATCTG	186
RESULT	3		
BH398899/c			
LOCUS	BH398899	687 bp	DNA linear GSS 11-DEC-2001
DEFINITION	AG-ND-159J17.TF ND-TAM Anopheles gambiae genomic clone AG-ND-159J17		
			. DNA sequence.

VERSION	00338699.1	GI.117493113
KEYWORDS	GSS.	
SOURCE	African malaria mosquito.	
ORGANISM	Anopheles gambiae	
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
	Phylogota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae	
	1 (bases 1 to 687)	
REFERENCE	1 (bases 1 to 687)	

unpublished (2001).
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Tel: 301 838 3543
Fax: 301 838 3543
Email: b1loftus@tigr.org

(TIGR). The BAC library was generated from *A. gambae* PEST strain DNA. All DNA was extracted from newly hatched instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.

Seq primer: M13 For
Class: BAC ends.
Location/Qualifiers
1. .687

FEATURES
source

```

/db_xref="taxon:7165"
/clone="AG-ND-159J17"
/clone_lib="ND-TAM"
/note="Vector: pEC3Acl; Site_1: HindIII"
BASE COUNT      188 a      146 c      106 g      247 t
ORIGIN

Query Match      63.9%; Score 19.8; DB 12; Length 687;
Best Local Similarity 77.4%; Pred. No. 3.4e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

D	b	221	CGGAACATATCAGCTAGCTGTAAACAATCTG	191
		RESULT	4	
		BF786352/c		
LOCUS		BF786352	820 bp	linear
DEFINITION		602112868F1 NCI_CGAP_Kid14 Mus musculus cdna clone IMAGE:4240961		
		5', mRNA sequence.		
ACCESSION		BF786352		
VERSION		BF786352.1	GI:12091388	
KEYWORDS		EST.		
SOURCE		house mouse.		
ORGANISM		Mus musculus		


```

/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      200 a 197 c 213 g 166 t
ORIGIN

Query Match      62.6%; Score 19.4; DB 10; Length 776;
Best Local Similarity 79.3%; Pred. No. 5.1e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 gtacctatcaggtgagcgtgagcaatttg 31
||||| ||||| || | ||||| |||
Db 132 GTACCTGTGCAGATACCCCTTGAGCAACTTG 104

RESULT 10
BE534304      646 bp mRNA linear EST 09-AUG-2000
LOCUS        601235365f1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3598768 5',
DEFINITION   mRNA sequence.
ACCESSION    BE534304
VERSION      BE534304.1 GI:9762949
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM8780 row: b column: 17
High quality sequence stop: 549.
Location/Qualifiers
1. .646
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3598768"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stages="5 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      158 a 144 c 188 g 156 t
ORIGIN

Query Match      61.9%; Score 19.2; DB 10; Length 646;
Best Local Similarity 87.5%; Pred. No. 5.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 ggtacctatcaggtgagcgtgagc 25
||||| ||||| ||| |||||
Db 59 GGTACCTATCAGATAGTCTGAGC 82

RESULT 11
BB852642/c 316 bp mRNA linear EST 26-NOV-2001
LOCUS        BB852642 RIKEN full-length enriched, B16 F10Y cells Mus musculus
DEFINITION   cDNA clone G370008H06 5', mRNA sequence.
ACCESSION    BB852642
VERSION      BB852642.1 GI:17094096
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        1 (bases 1 to 316)
JOURNAL      Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
COMMENT      Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .316
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G370008H06"
/clone_lib="RIKEN full-length enriched, B16 F10Y cells"
/cell_type="B16 F10Y cells"
BASE COUNT      69 a 89 c 81 g 77 t
ORIGIN

Query Match      61.3%; Score 19; DB 9; Length 316;
Best Local Similarity 81.5%; Pred. No. 6.1e+02;

```

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cgtaccatcaggtagcgtgagca 27
||||| ||||| ||||| ||||| |||||
Db 32 CGGACCTCCAGCGCGCTGGCAA 6

RESULT 12

AU005717/c

LOCUS

DEFINITION AU005717 Bombyx mori p50(Daizo) Bombyx mori cDNA clone wv40011, EST 19-JAN-1999

ACCESSION

VERSION AU005717

KEYWORDS

SOURCE

ORGANISM

Bombyx mori

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia

; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 361)

AUTHORS Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.

TITLE Establishment of cDNA database of Bombyx mori

JOURNAL Unpublished (1999)

COMMENT Contact: Mita K

Genome Research Group

National Institute of Radiological Sciences

Anagawa 4-9-1, Inage, Chiba 263-8555, Japan

Email: kmita@nirs.go.jp

PROJECT = 'CREST project by JST'.

FEATURES

source

1..361

Location/Qualifiers

/organism="Bombyx mori"

/strain="p50(Daizo)"

/db_xref="taxon:7091"

/clone="wv40011"

/clone_lib="Bombyx mori p50(Daizo)"

BASE COUNT 105 a 98 c 81 g 77 t

ORIGIN

Query Match 61.3%; Score 19; DB 9; Length 361;

Best Local Similarity 81.5%; Pred. No. 6.3e+02;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ggtaccatcaggtagcgtgagcaat 28

||||| ||||| ||||| ||||| |||||

Db 98 GGTGCCGATCACATAGCGCTGAGCAAT 72

RESULT 13

BG228080

LOCUS

DEFINITION BG228080 x1 Soares NMAX_maxillary process Mus musculus cDNA clone IMAGE:3513271 3' similar to SW:RBBI_HUMAN P29374 RETINOBLASTOMA BINDING PROTEIN 1 ; mRNA sequence.

ACCESSION

VERSION BG228080

KEYWORDS

SOURCE

house mouse

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 379)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1393847

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.

FEATURES

source

1..379

Location/Qualifiers

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:3513271"

/clone_lib="Soares_NMAX_maxillary_process"

/tissue_type="maxillary process"

/lab_host="DH10B (phage-resistant)"

/note="vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTT 3'] ,

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 116 a 67 c 98 g 98 t

ORIGIN

Query Match 61.3%; Score 19; DB 10; Length 379;

Best Local Similarity 81.5%; Pred. No. 6.4e+02;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ggtaccatcaggtagcgtgagcaat 28

||||| ||||| ||||| ||||| |||||

Db 203 GGTGCCATCAGGAGCGGTATCAAT 229

RESULT 14

BE634195/c

LOCUS

DEFINITION BE634195 y1 Soares mouse 3NBMS Mus musculus cDNA clone IMAGE:3412493 5' similar to gb:S66427 RETINOBLASTOMA BINDING PROTEIN 1 (HUMAN); , mRNA sequence.

ACCESSION

VERSION BE634195

KEYWORDS

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 390)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1088305

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..390

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:3412493"

/clone_lib="Soares mouse 3NBMS"

/sex="male"

/tissue_type="Spleen"

/dev_stage="4 weeks"

/lab_host="DH10B"

/note="vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCTTTTTTTTTTTT 3'] ; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7/3 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 93 a 102 c 73 g 122 t
ORIGIN

Query Match 61.3%; Score 19; DB 10; Length 390;

Best Local Similarity 81.5%; Pred. No. 6.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtacgctgagcaat 28
||| ||||||||| ||||| |||||
Db 155 GGTGCTATCAGGAAGCGGTTATCAAT 129

RESULT 15

BE986391

LOCUS

BE986391 390 bp mRNA linear EST 05-OCT-2000
UI-M-CG0p-bgr-h-10-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
UI-M-CG0p-bgr-h-10-0-UI 3', mRNA sequence.

ACCESSION

BE986391

VERSION

BE986391.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA-No.

FEATURES

source

Location/Qualifiers

1. .390
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0p-bgr-h-10-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="vector: pMT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"

BASE COUNT 115 a 74 c 109 g 91 t 1 others
ORIGIN

Query Match

61.3%; Score 19; DB 10; Length 390;

Best Local Similarity 81.5%; Pred. No. 6.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtacgctgagcaat 28
||| ||||||||| ||||| |||||
Db 282 GGTGCTATCAGGAAGCGGTTATCAAT 308

Search completed: August 31, 2002, 15:24:39
Job time: 6382 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:27:36 ; Search time 153.1 Seconds
(without alignments)
49,736 Million cell updates/sec

Title: US-09-810-861b-2
Perfect score: 31
Sequence: 1 cggtaacctatcaggtagcgtgagcaatttg 31

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.8	60.6	16075	3	US-09-096-942-1
2	18.8	60.6	16075	3	US-09-096-867-1
c 3	18.2	58.7	829	4	US-08-998-416-495
4	18.2	58.7	1366	3	US-09-191-099-4
c 5	18	58.1	3016	2	US-08-318-826A-7
c 6	18	58.1	3016	2	US-08-370-156-5
c 7	18	58.1	3016	3	US-08-814-095-5
c 8	17.8	57.4	3096	2	US-08-318-826A-6
c 9	17.8	57.4	3096	2	US-08-370-156-3
c 10	17.8	57.4	3096	3	US-08-814-095-3
c 11	17.8	57.4	3096	3	US-08-814-095-7
c 12	17.4	56.1	1845	1	US-07-732-962A-1
c 13	17.4	56.1	1845	5	PCT-US92-06106-1
c 14	17.4	56.1	2256	2	US-08-318-826A-5
c 15	17.4	56.1	2256	2	US-08-370-156-1
c 16	17.4	56.1	2256	3	US-08-814-095-1
17	17.4	56.1	3499	4	US-08-857-076-43
18	17.2	55.5	1462	3	US-09-191-099-3
c 19	17.2	55.5	2133	3	US-08-808-032-1
20	17	54.8	3955	1	US-08-229-515A-14
21	17	54.8	3955	1	US-08-645-865-14
22	17	54.8	9045	3	US-09-121-321-1
23	17	54.8	9045	4	US-08-933-803A-1
24	17	54.8	50937	4	US-09-428-517-1
25	16.8	54.2	1325	2	US-08-632-470-50
c 26	16.6	53.5	595	6	5242807-1
27	16.6	53.5	1413	3	US-09-191-099-1

```

28 16.6 53.5 1503 4 US-08-943-571-1
c 29 16.4 52.9 9785 1 US-08-319-387-1
c 30 16.2 52.3 861 2 US-08-743-637B-166
c 31 16.2 52.3 861 3 US-08-526-840B-166
c 32 16.2 52.3 1920 1 US-08-746-789A-1
c 33 16.2 52.3 3000 2 US-08-928-692-9
c 34 16.2 52.3 3060 2 US-08-560-398-1
c 35 16 51.6 50 4 US-09-390-867A-28
c 36 16 51.6 50 4 US-09-548-260-28
c 37 16 51.6 1038 4 US-09-004-838-128
c 38 16 51.6 3494 4 US-09-139-802-200
c 39 16 51.6 3645 2 US-08-663-112-1
c 40 16 51.6 5427 3 US-09-009-913-2
c 41 16 51.6 5510 3 US-09-009-913-3
c 42 16 51.6 5667 3 US-09-009-913-4
c 43 15.8 51.0 29 1 US-08-222-177A-460
c 44 15.8 51.0 1224 2 US-08-210-762E-19
c 45 15.8 51.0 1224 4 US-09-106-075A-18

```

ALIGNMENTS

```

RESULT 1
US-09-096-942-1
; Sequence 1, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,942
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-942-1

```

Query Match 60.6% Score 18.8; DB 3; Length 16075;
Best Local Similarity 90.9%; Pred. No. 5.7;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 10 tcaggtagcgtgagcaatttg 31
|||||
DB 11351 tcaggtagcgtgagcgtgtg 11372

```

```

RESULT 2
US-09-096-867-1
; Sequence 1, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867

```

; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-867-1

Query Match 60.6%; Score 18.8; DB 3; Length 16075;
Best Local Similarity 90.9%; Pred. No. 5.7;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 tcaggtagcgctgagcaatttg 31
||||| ||||| ||||| ||||| |||||

Db 11351 tcaggtagcgctgagcgatgtg 11372

RESULT 3

US-08-998-416-495/C
; Sequence 495, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Melgs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8367
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 495:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 829 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1350RP
US-08-998-416-495

Query Match 58.7%; Score 18.2; DB 4; Length 829;
Best Local Similarity 83.3%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 ctatcaggtagcgctgagcaattt 30
||| ||||| ||||| ||||| |||||

Db 826 CCAAGAGGTAGNGCTGAGCAATT 803

RESULT 4

US-09-191-099-4
; Sequence 4, Application US/09191099
; Patent No. 6096323
; GENERAL INFORMATION:
; APPLICANT: Walker, Richard L.
; APPLICANT: Read, Deryck H.
; APPLICANT: Hird, David W.
; APPLICANT: Lefebvre, Rance B.
; APPLICANT: Berry, Steven L.
; APPLICANT: Cullor, James S.
; APPLICANT: Lefler, Hank M.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Vaccine Against Papillomatous Digital Dermatitis (PDD)
; FILE REFERENCE: 023070-081110DS
; CURRENT APPLICATION NUMBER: US/09/191,099
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: US 08/943,571
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Treponema sp.
; FEATURE:
; OTHER INFORMATION: 16S rRNA for spirochete 9-3301
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (78)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (334)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1006)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1121)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1200)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1228)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1250)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1262)
; OTHER INFORMATION: n = unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1271)

US-08-318-B26A-1

RESULT 7
US-08-814-095-5/c
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: zakut, HaIm

APPLICANT: Shani, Moshe
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 No. 602518thwestern Highway, Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3016 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Alternatively spliced AChE
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 160..2010
US-08-814-095-5

Query Match 58.1%; Score 18; DB 3; Length 3016;
Best Local Similarity 80.8%; Pred. No. 9.7;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 cctatcaggtagcgtgagcaatttg 31
|| | ||| ||||| ||||| |||||
DB 1888 CCGAGCGGTGGCGTGAGCAATTG 1863

RESULT 8
US-08-318-826A-6/c
Sequence 6, Application US/08318826A
Patent No. 5891725
GENERAL INFORMATION:
APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Eckstein, Fritz
TITLE OF INVENTION: Synthetic Antisense
TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
TITLE OF INVENTION: Containing Them
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 5891725thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,826A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 2391.00001
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3096 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 160..1959
OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
OTHER INFORMATION: 3, 4, 5 and the translated portion of Intron 4 (readthrough
US-08-318-826A-6

Query Match 57.4%; Score 17.8; DB 2; Length 3096;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 caggtagcgtgagcaatttg 31
||| ||| ||||| ||||| |||||
DB 1883 CCGTGGCGCTGAGCAATTG 1863

RESULT 9
US-08-370-156-3/c
Sequence 3, Application US/08370156
Patent No. 5932780
GENERAL INFORMATION:
APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Shani, Moshe
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reising, Ethlington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,156
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955

; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
; US-08-370-156-3

Query Match 57.4%; Score 17.8; DB 2; Length 3096;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 11 caggtagcgcctgagcaattg 31
| ||| ||||| ||||| |||||
Db 1883 CCGGTGGCGCTGAGCAATTG 1863

RESULT 10
US-08-814-095-3/c
; Sequence 3, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced ACHE
; DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
; DESCRIPTION: of intron 4 (readthrough)"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

; LOCATION: 160..1959
; US-08-814-095-3

Query Match 57.4%; Score 17.8; DB 3; Length 3096;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 11 caggtagcgcctgagcaattg 31
| ||| ||||| ||||| |||||
Db 1883 CCGGTGGCGCTGAGCAATTG 1863

RESULT 11
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE
; DESCRIPTION: promoter, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,962A
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
; US-07-732-962A-1

Query Match 56.1%; Score 17.4; DB 1; Length 1845;
Best Local Similarity 94.7%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtagcgtgagcaatttg 31
||| ||||| ||||| |||||
Db 1722 GGTGGCGCTGAGCAATTG 1704

RESULT 13
PCT-US92-06106-1/c
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Melr
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06106
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
; PCT-US92-06106-1

Query Match 56.1%; Score 17.4; DB 5; Length 1845;
Best Local Similarity 94.7%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtagcgtgagcaatttg 31
||| ||||| ||||| |||||
Db 1722 GGTGGCGCTGAGCAATTG 1704

RESULT 14
US-08-318-826A-5/c
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens

```
;
; FEATURE:
; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match          56.1%; Score 17.4; DB 2; Length 2256;
Best Local Similarity 94.7%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtgcgcctgagcaattg 31
   ||| ||||| ||||| |||
Db 1881 GGTGGCGCTGAGCAATTG 1863

RESULT 15
US-08-370-156-1/c
; Sequence 1, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soteg, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethlington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-156-1
```

```
Query Match          56.1%; Score 17.4; DB 2; Length 2256;
Best Local Similarity 94.7%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtgcgcctgagcaattg 31
   ||| ||||| ||||| |||
Db 1881 GGTGGCGCTGAGCAATTG 1863
```

Search completed: August 31, 2002, 15:27:40
Job time: 6422 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:36:11 ; Search time 9324.54 Seconds
(without alignments)
65.083 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatatctgcagccatggctagggccccgc 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
--------	-----	-------	-------	--------	----	----	-------------

ALIGNMENTS

RESULT 1

AX275252

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .29

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="primer pAChE-Nco, derived from human AChE gene and

AX275252 Sequence

AF432517 Avena sat

AF432531 Hordeum v

AF432532 Hordeum v

AC010395 Homo sapi

AC012610 Homo sapi

AX018091 Sequence

BC006151 Homo sapi

AX014146 Sequence

AX018186 Sequence

AC010325 Homo sapi

AC027602 Homo sapi

AB026686 Physcomit

M95404 Callus gall

AC004206 Giardia i

AC052537 Giardia i

AC049241 Giardia i

AF054819 Mus muscu

D86386 Catharanthu

AJ277867 Pimephale

BC003590 Homo sapi

I09460 Sequence 5

AK024433 Homo sapi

J03252 Human alk

D50923 Human mRNA

AE006947 Mycobacte

AC014904 Drosophil

284724 Mycobacteri

AC015005 Drosophil

AD000014 Mycobacte

AC090716 Homo sapi

AC098681 Homo sapi

AP001435 Homo sapi

AC105820 Rattus no

AC098662 Rattus no

AP000008 Homo sapi

AP000149 Homo sapi

AC012387 Drosophil

AP003787 Homo sapi

Continuation (3 of

AC008615 Homo sapi

AP000704 Homo sapi

AC103965 Homo sapi

AL603805 Rattus no

AL356220 Human DNA

AX275252 Sequence 1 from Patent WO0171014.

AX275252

AX275252.1 GI:16547672

synthetic construct.

synthetic construct

artificial sequence.

1 (sites)

Mor,T., Soreq,H., Arntzen,H. and Mason,H.

Expression of recombinant human acetylcholinesterase in transgenic

plants

Patent: WO 0171014-A 1 27-SEP-2001;

BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,

Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,

Hugh S. (US)

Location/Qualifiers

1. .29

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="primer pAChE-Nco, derived from human AChE gene and

```

BASE COUNT      5 a      11 c      8 g      5 t
ORIGIN
modified to introduce an Nco I restriction site"

Query Match      100.0%; Score 29; DB 6; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatattgcagccatggctaggcccccgc 29
|||||
Db 1 GATATCTCAGCCATGGCTAGGCCCCCGC 29

RESULT 2
AF432517      474 bp      DNA      linear      PLN 30-JAN-2002
LOCUS
DEFINITION      Avena sativa clone As3431 transposon mariner-like transposase
ACCESSION      AF432517
VERSION      AF432517.1 GI:18419466
KEYWORDS      .
SOURCE      oat.
ORGANISM      Avena sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Poaceae; Avena.
REFERENCE      1 (bases 1 to 474)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Mariner-like transposases are widespread and diverse in flowering
Plants
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE      21642679
PUBMED      11756687
REFERENCE      2 (bases 1 to 474)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Direct Submission
JOURNAL      Submitted (02-OCT-2001) Botany, University of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source
1. .474
/organism="Avena sativa"
/db_xref="taxon:4498"
/clone="As3431"
<1. .>474
/gene="transposase"
/pseudo
repeat_region      <1. .>474
BASE COUNT      161 a      73 c      120 g      120 t
ORIGIN

Query Match      62.1%; Score 18; DB 8; Length 474;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 tgcagccatggctaggcc 24
|||||
Db 120 TGCAGCCATGGCTAGGCC 137

RESULT 3
AF432531      485 bp      DNA      linear      PLN 30-JAN-2002
LOCUS
DEFINITION      Hordeum vulgare clone Hv2936 transposon mariner-like transposase
ACCESSION      AF432531
VERSION      AF432531.1 GI:18419481
KEYWORDS      .
SOURCE      barley.
ORGANISM      Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticaceae; Hordeum.
REFERENCE      1 (bases 1 to 485)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Mariner-like transposases are widespread and diverse in flowering
Plants
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE      21642679
PUBMED      11756687
REFERENCE      2 (bases 1 to 485)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Direct Submission
JOURNAL      Submitted (02-OCT-2001) Botany, University of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source
1. .485
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="Hv2936"
<1. .>485
/transposon="mariner-like"
/gene="transposase"
/pseudo
repeat_region
gene
BASE COUNT      166 a      76 c      116 g      127 t
ORIGIN

```

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticaceae; Hordeum.
REFERENCE      1 (bases 1 to 485)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Mariner-like transposases are widespread and diverse in flowering
Plants
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE      21642679
PUBMED      11756687
REFERENCE      2 (bases 1 to 485)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Direct Submission
JOURNAL      Submitted (02-OCT-2001) Botany, University of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source
1. .485
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="Hv2936"
<1. .>485
/transposon="mariner-like"
/gene="transposase"
/pseudo
repeat_region
gene
BASE COUNT      168 a      73 c      115 g      129 t
ORIGIN

Query Match      62.1%; Score 18; DB 8; Length 485;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 tgcagccatggctaggcc 24
|||||
Db 120 TGCAGCCATGGCTAGGCC 137

RESULT 4
AF432532      485 bp      DNA      linear      PLN 30-JAN-2002
LOCUS
DEFINITION      Hordeum vulgare clone Hv2939 transposon mariner-like transposase
ACCESSION      AF432532
VERSION      AF432532.1 GI:18419482
KEYWORDS      .
SOURCE      barley.
ORGANISM      Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticaceae; Hordeum.
REFERENCE      1 (bases 1 to 485)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Mariner-like transposases are widespread and diverse in flowering
Plants
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE      21642679
PUBMED      11756687
REFERENCE      2 (bases 1 to 485)
AUTHORS      Feschotte,C. and Wessler,S.R.
TITLE      Direct Submission
JOURNAL      Submitted (02-OCT-2001) Botany, University of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source
1. .485
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="Hv2939"
<1. .>485
/transposon="mariner-like"
/gene="transposase"
/pseudo
repeat_region
gene
BASE COUNT      166 a      76 c      116 g      127 t
ORIGIN

```

```

ORIGIN
Query Match      62.1%; Score 18; DB 8; Length 485;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 tgcagccatgctaggcc 24
Db 120 TGCAGCCATGCTAGGCC 137

RESULT 5
AC010395 AC010395 130615 bp DNA linear PRI 29-SEP-2000
LOCUS Homo sapiens chromosome 5 clone CTD-2089024, complete sequence.
DEFINITION AC010395
ACCESSION AC010395.6 GI:10337627
VERSION HTG.
KEYWORDS SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 130615)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 130615)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT On Sep 29, 2000 this sequence version replaced gi:7711435.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.7.
FEATURES             source
    Location/Qualifiers
        .1..130615
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="5"
            /clone="CTD-2089024"
BASE COUNT   39938 a 24537 c 25833 g 40307 t
ORIGIN

Query Match      62.1%; Score 18; DB 9; Length 130615;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctcagccatggctagg 22
Db 50672 TCTCAGCCATGGCTAGG 50689

RESULT 6
AC012610 AC012610 135248 bp DNA linear PRI 15-AUG-2001
LOCUS Homo sapiens chromosome 5 clone CTC-540012, complete sequence.
DEFINITION AC012610
ACCESSION AC012610.6 GI:15187228
VERSION HTG.
KEYWORDS SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135248)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmänn,B., Rosenthal,A., Pillarsky,C.
TITLE Human nucleic acid sequences from prostate tumour tissue
JOURNAL Patent: WO 9946374-A 48 16-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUNG (DE); PILLARSKY CHRISTIAN (DE)
REFERENCE 2 (bases 1 to 135248)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT On Aug 15, 2001 this sequence version replaced gi:123314599.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.4.
FEATURES             source
    Location/Qualifiers
        .1..135248
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="5"
            /clone="CTC-540012"
BASE COUNT   41508 a 25271 c 26886 g 41583 t
ORIGIN

Query Match      62.1%; Score 18; DB 9; Length 135248;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctcgaccatggctagg 22
Db 55313 TCTCGACCATGGCTAGG 55330

RESULT 7
AX018091 AX018091 906 bp DNA linear PAT 07
LOCUS Sequence 48 from Patent W09946374.
DEFINITION AX018091
ACCESSION AX018091
VERSION AX018091.1 GI:10042542
KEYWORDS SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmänn,B., Rosenthal,A., Pillarsky,C.
TITLE Human nucleic acid sequences from prostate tumour tissue
JOURNAL Patent: WO 9946374-A 48 16-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUNG (DE); PILLARSKY CHRISTIAN (DE)
REFERENCE 2 (bases 1 to 135248)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT On Aug 15, 2001 this sequence version replaced gi:123314599.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.4.
FEATURES             source
    Location/Qualifiers
        .1..135248
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="5"
            /clone="CTC-540012"
BASE COUNT   41508 a 25271 c 26886 g 41583 t
ORIGIN

```

Query Match 58.6%; Score 17; DB 6; Length 906;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 ctgcagccatgctagg 22
 |||||
 Db 228 CTGCAGCCATGCTAGG 244

RESULT 8
 LOCUS BC006151 1316 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, clone MGC:13170 IMAGE:3354369, mRNA, complete cds.
 ACCESSION BC006151
 VERSION BC006151.1 GI:13544031
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1316)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 18 Row: d Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES Location/Qualifiers
 source 1..1316
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:13170 IMAGE:3354369"
 /tissue_type="Eye, retinoblastoma"
 /clone_lib="NIH_MGC_16"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 CDS 555..908
 /codon_start=1
 /product="Unknown (protein for MGC:13170)"
 /protein_id="AAH06151.1"
 /db_xref="GI:13544032"
 /translation="MTVLEAVLEIQAITGSRLLSWPGPAPPGSCHDPTQCTWLL
 SHTPRRWISGLPRASRLGEEPPPLPCQAYGEELSIKRRETAWLSRDTWPNGA
 PGVKQRLIGLELLV"
 BASE COUNT 269 a 419 c 387 g 241 t
 ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 1316;

Best Local Similarity 100.0%; Pred. No. 28;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 ctgcagccatgctagg 22
 |||||
 Db 85 CTGCAGCCATGCTAGG 101

RESULT 9
 LOCUS AX014146 1467 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 11 from Patent WO9954447.
 ACCESSION AX014146
 VERSION AX014146.1 GI:10040593
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1467)
 AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and
 Pilarsky, C.
 TITLE Human nucleic acid sequences of bladder tumour tissue
 JOURNAL Patent: WO 9954447-A 11 28-OCT-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
 (DE); PILARSKY CHRISTIAN (DE)

FEATURES Location/Qualifiers
 source 1..1467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 307 a 463 c 432 g 265 t
 ORIGIN

Query Match 58.6%; Score 17; DB 6; Length 1467;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 ctgcagccatgctagg 22
 |||||
 Db 229 CTGCAGCCATGCTAGG 245

RESULT 10
 LOCUS AX018186 1467 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 159 from Patent WO9946374.
 ACCESSION AX018186
 VERSION AX018186.1 GI:10042564
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1467)
 AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and
 Pilarsky, C.
 TITLE Human nucleic acid sequences from prostate tumour tissue
 JOURNAL Patent: WO 9946374-A 159 16-SEP-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
 (DE); PILARSKY CHRISTIAN (DE)

FEATURES Location/Qualifiers
 source 1..1467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 307 a 463 c 432 g 265 t
 ORIGIN

Query Match 58.6%; Score 17; DB 6; Length 1467;
 Best Local Similarity 100.0%; Pred. No. 27;

* consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 8149: contig of 8149 bp in length
 8150 8249: gap of 100 bp
 8250 9592: contig of 1343 bp in length
 9593 9692: gap of 100 bp
 9693 10733: contig of 1041 bp in length
 10734 10833: gap of 100 bp
 10834 13519: contig of 2686 bp in length
 13520 13619: gap of 100 bp
 13620 17510: contig of 3891 bp in length
 17511 17610: gap of 100 bp
 17611 24602: contig of 6992 bp in length
 24603 24702: gap of 100 bp
 24703 35434: contig of 10732 bp in length
 35435 35534: gap of 100 bp
 35535 124474: contig of 88940 bp in length
 124475 124574: gap of 100 bp
 124575 134664: contig of 10090 bp in length
 134665 134764: gap of 100 bp
 134765 162343: contig of 27579 bp in length
 162344 162443: gap of 100 bp
 162444 208917: contig of 46474 bp in length
 208918 209017: gap of 100 bp
 209018 217346: contig of 8329 bp in length.

FEATURES

source
 1. .217346
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-795B6"
 /clone_lib="RPC1-11 Human Male BAC"
 1. .8149
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 8250. .9592
 /note="assembly_fragment"
 9693. .10733
 /note="assembly_fragment"
 10834. .13519
 /note="assembly_fragment"
 13620. .17510
 /note="assembly_fragment"
 17611. .24602
 /note="assembly_fragment"
 24703. .35434
 /note="assembly_fragment"
 35535. .124474
 /note="assembly_fragment"
 124575. .134664
 /note="assembly_fragment"
 134765. .162343
 /note="assembly_fragment"
 162444. .208917
 /note="assembly_fragment"
 209018. .217346
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"
 55109 a 54154 c 53923 g 53053 t 1107 others

Query Match 58.6%; Score 17; DB 2; Length 217346;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatgctagg 22
 |||||
 Db 6135 CTGCAGCCATGCTAGG 6119

RESULT 13

AB026686
 LOCUS Physcomitrella patens mRNA for chlorophyll a/b-binding protein
 DEFINITION precursor, complete cds.
 ACCESSION AB026686
 VERSION AB026686.1 GI:4757370
 KEYWORDS chlorophyll a/b-binding protein precursor.
 SOURCE Physcomitrella patens cDNA to mRNA.
 ORGANISM Physcomitrella patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (sites)
 AUTHORS Kiyosue,T. and Wada,M.
 TITLE Three cDNA sequences for light inducible genes from Physcomitrella
 patens
 JOURNAL Unpublished (1999)
 REFERENCE 2 (bases 1 to 964)
 AUTHORS Kiyosue,T.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-1999) Tomohiro Kiyosue, National Institute for
 Basic Biology (NIBB), Biological Regulation Division; Myoudaigai,
 Okazaki, Aichi 444-8585, Japan (E-mail:tkiyosue@nibb.ac.jp,
 Tel:81-564-55-7610, Fax:81-564-55-7611)

FEATURES

source
 1. .964
 /organism="Physcomitrella patens"
 /db_xref="taxon:3218"
 40. .846
 /codon_start=1
 /product="chlorophyll a/b-binding protein precursor"
 /protein_id="BAA77273.1"
 /db_xref="GI:4757371"
 /translation="MATAATAMNSTVLQAGALLKPVSELSRKVNAGEARVTRKTVSK
 SSGSDSIWYGADRPKFLGPGSGTSLNGEAGYDGTAGLSSDPETFAFNRELEV
 ITHRWAMGALGCLPELLAKSGVGEAVWFKAGQIFSEGGLDYLGNPSLVHAQSI
 LAIWACQVILMGAVGEYRVAGGPDVDTPIYGGSFDPGLGADDPTDTAEKLVKEIK
 NGRLAMFSMGFFVQAIVTGKPLNDLHLPADVANNANWAYATNFVPGN"
 BASE COUNT 206 a 280 c 262 g 216 t
 ORIGIN

Query Match 55.2%; Score 16; DB 8; Length 964;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatggc 18
 |||||
 Db 29 TATCTGCAGCCATGGC 44

RESULT 14

CHKPRION
 LOCUS Physcomitrella patens mRNA for chlorophyll a/b-binding protein
 DEFINITION precursor, complete cds.
 ACCESSION M95404
 VERSION M95404.1 GI:212610
 KEYWORDS chicken.
 SOURCE Gallus gallus
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 969)
 AUTHORS Gabriel,J.-M., Oesch,B., Kretzschmar,H., Scott,M. and Prusiner,S.B.
 TITLE Molecular cloning of a candidate chicken prion protein
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

MEDLINE	33028411	Location/Qualifiers
FEATURES	source	
	1..969	
	/organism="Gallus gallus"	
	/strain="White Leghorn"	
	/sub_species="domesticus"	
	/db_xref="taxon:9031"	
	/sex="male"	
	/tissue_type="liver"	
	/dev_stage="adult"	
	/germline	
	/tissue_lib="Clontech"	
Intron	<1...34	
	/gene="Prn-p"	
	/number=1	
gene	<1..>969	
	/gene="Prn-p"	
exon	35..>969	
	/gene="Prn-p"	
	/number=2	
5'UTR	<35..36	
	/gene="Prn-p"	
mRNA	<35..>969	
	/gene="Prn-p"	
	/product="prion protein"	
CDS	37..858	
	/gene="Prn-p"	
	/codon_start=1	
	/product="prion protein"	
	/protein_id="AAC28970.1"	
	/db_xref="GI:212611"	
	/translation="MARLLTTCCLLALLAACTDVALSKKGKPGSGGAGSHRQPSYPGPGPHNPGYPHNPGYPHNPGYPHNPGYPHNPGYPHNPGYGVGQGVNPSGGSYHNOQPKPKPTNFHVGAGAAAGVGLGSGYAMGRVMSGMNTHFDSPEYRWSGNSARYPNRYRDYSSVPQDVADCFNITVTEYSIGPAKKNTSEAVAANQTEVEHEKNVTKVIREMCVQYREYRLASGIQHPADTWLAVLLLLLTFLFAWH"	
slg_peptide	37..108	
	/gene="Prn-p"	
mat_peptide	109..855	
	/gene="Prn-p"	
	/product="prion protein"	
3'UTR	859..>969	
	/gene="Prn-p"	
BASE COUNT	194 a 331 c 270 g 174 t	
ORIGIN		
	Query Match 55.2%; Score 16; DB 5; Length 969;	
	Best Local Similarity 100.0%; Pred. No. 1.1e+02;	
	Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	8 gcagccatgctagc 23	
DB	31 GCACCATGCGCTAGC 46	
RESULT	15	
AC064206/C	standard; DNA; HTG; 1002 BP.	
XX	AC064206	
XX	AC064206;	
XX	AC064206.1	
XX		
DT	26-APR-2000 (Rel. 63, Created)	
DT	26-APR-2000 (Rel. 63, Last updated, Version 1)	
XX		
DE	Giardia intestinalis clone NJ3676 strain WB-C6, LOW-PASS SEQUENCE SAMPLING.	
XX	HTG; HTGS_PHASE0.	
XX		
OS	Giardia intestinalis	
OC	Eukaryota; Diplomonadida; Hexamitidae; Giardia.	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:18:19 ; Search time 824.51 Seconds
(Without alignments)
60.388 Million cell updates/sec

Title: US-09-810-861B-1

Perfect score: 29

Sequence: 1 gatatctgcagcatgctagggccccgc 29

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-nbml/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	29	AAS17545	Human acetylcholin
c	2	18	62.1	3023 23 AAS77493	DNA encoding novel
c	3	18	62.1	3684 23 AAS80007	DNA encoding novel
4	17	58.6	271 21 AAC00220	Human secreted pro	
5	17	58.6	906 20 AAZ33470	Human prostate can	
6	17	58.6	1452 22 AAI89867	Human polynucleoti	
7	17	58.6	1466 21 AAF16096	Human prostate can	
8	17	58.6	1467 20 AAZ24399	Human bladder tumo	
9	16	55.2	473 21 AAA79626	Pinus radiata cell	

c	10	16	55.2	4356	21	AA79736	Pinus radiata cell
c	11	16	55.2	4481	10	AAN90958	Sequence encoding
c	12	16	55.2	4487	13	AAQ20514	Encodes germ cell
c	13	15	51.7	51	22	AAL31467	Human SNP oligonuc
c	14	15	51.7	204	19	AA110671	Human biallelic po
c	15	15	51.7	375	22	AAF66350	Novel human polynu
c	16	15	51.7	517	22	ABA61689	Human foetal liver
c	17	15	51.7	517	22	ABA29336	Probe #7802 for ge
c	18	15	51.7	517	22	AAK09991	Human brain expres
c	19	15	51.7	517	22	AAK35885	Human bone marrow
c	20	15	51.7	517	22	AA117169	Probe #10287 used
c	21	15	51.7	517	22	AA141601	Human AFP protein
c	22	15	51.7	657	22	AAH52128	DNA encoding a pol
c	23	15	51.7	685	21	AA63739	Arabidopsis thalia
c	24	15	51.7	1215	21	AAC46923	Mouse ischaemic co
c	25	15	51.7	1360	24	AB199523	DNA encoding novel
c	26	15	51.7	1921	23	AA576387	DNA encoding a pol
c	27	15	51.7	1994	21	AA63740	DNA encoding molec
c	28	15	51.7	2109	22	AA02049	CDNA encoding nove
c	29	15	51.7	2390	22	AA545000	Drosophila melanog
c	30	15	51.7	2439	23	ABL05705	Human polynucleoti
c	31	15	51.7	6454	22	AA157932	Human polynucleoti
c	32	15	51.7	6742	22	AA159718	Human polynucleoti
c	33	15	51.7	12143	23	ABL05704	Drosophila melanog
c	34	15	51.7	12807	22	AA105769	Human reproductive
c	35	15	51.7	13165	22	ABA07140	Human pancreatic c
c	36	15	51.7	13165	22	AAK89348	Human digestive sy
c	37	15	51.7	16831	23	AA559607	Propionibacterium
c	38	15	51.7	42488	22	AAK66772	Human immune/haema
c	39	14	48.3	22	13	AAQ27953	CAS5 primer 2. Sy
c	40	14	48.3	24	13	AAQ35538	Spacer MRSYN20. S
c	41	14	48.3	24	13	AAQ35539	Spacer MRSYN19. S
c	42	14	48.3	36	21	AA73841	Human bone marrow
c	43	14	48.3	107	18	AA73876	Maize G9 core prom
c	44	14	48.3	107	20	AA227552	G9 core promoter s
c	45	14	48.3	140	11	AAQ06660	Synthetic antifree

ALIGNMENTS

RESULT 1
AAS17545
ID AAS17545 standard; DNA; 29 BP.

XX AAS17545;

XX 25-FEB-2002 (first entry)

DE Human acetylcholinesterase, PCR primer AChE-Nco.

XX Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ss.

OS Homo sapiens.

OS Synthetic.

XX WO200171014-A2.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08468.

XX 17-MAR-2000; 2000US-190440P.

XX (MORT/) MOR T.

XX (SORE/) SOREQ H.

XX (ARNT/) ARNTZEN C.

XX (MASO/) MASON H.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

```
PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
XX Claim 11; Page 4; 42pp; English.
XX
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of acetylcholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents PCR primer AChE-Nco used to amplify
CC human AChE as described in the method of the invention.
XX
XX Sequence 29 BP; 5 A; 11 C; 8 G; 5 T; 0 other;
SQ

Query Match 100.0%; Score 29; DB 24; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatattcgagccatggctagggccgcg 29
Db 1 gatattcgagccatggctagggccgcg 29

RESULT 2
AAS77493/C
ID AAS77493 standard; cDNA; 3023 BP.
XX
XX AAS77493;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #13297.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
```

```
XX WPI; 2001-639362/73.
DR P-PSDB; ABG13306.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID NO 13297; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3023 BP; 1083 A; 698 C; 631 G; 611 T; 0 other;
SQ

Query Match 62.1%; Score 18; DB 23; Length 3023;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagg 22
Db 1362 TCtgcagccatggctagg 1345

RESULT 3
AAS80007/C
ID AAS80007 standard; cDNA; 3684 BP.
XX
XX AAS80007;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #15811.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
```

DR WPI: 2001-639362/73.
 DR P-PSDB; ABG15820.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 1: SEQ ID No 15811; 103pp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3684 BP; 1391 A; 856 C; 712 G; 725 T; 0 other;

Query Match 62.1%; Score 18; DB 23; Length 3684;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ctgcagccatgctagg 22
 |||||
 Db 2628 TCTGCACCATGCTAGG 2611

RESULT 4
 AAC00220
 ID AAC00220 standard; cDNA; 271 BP.
 XX
 AC AAC00220;
 XX
 DT 06-OCT-2000 (first entry)
 DE Human secreted protein 5' EST, SEQ ID NO: 218.
 KW Human: 5' EST: expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX Homo sapiens.
 XX EP1033401-A2.
 XX 06-SEP-2000.
 XX 21-FEB-2000; 2000EP-0200610.
 XX 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Giordano J;

PI WPI: 2000-500381/45.
 DR P-PSDB; AAG00214.
 DR

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1: SEQ ID 218; 71pp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX
 SQ Sequence 271 BP; 53 A; 84 C; 76 G; 56 T; 2 other;

Query Match 58.6%; Score 17; DB 21; Length 271;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatgctagg 22
 |||||
 Db 200 ctgcagccatgctagg 216

RESULT 5
 AAZ33470
 ID AAZ33470 standard; cDNA; 906 BP.
 XX
 AC AAZ33470;
 XX
 DT 08-DEC-1999 (first entry)
 DE Human prostate cancer-associated EST 48.
 XX
 XX Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
 KW gene therapy; tissue specificity human; ss.
 XX Homo sapiens.
 XX DE19011193-A1.
 XX 16-SEP-1999.
 XX 10-MAR-1998; 98DE-1011193.
 XX 10-MAR-1998; 98DE-1011193.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E, Rosenthal A;

PI WPI: 1999-519628/44.
 DR P-PSDB; AAY48276, AAY48277, AAY48278, AAY48279, AAY48280, AAY48281,
 DR AAY48282.
 XX New nucleic acid expressed at high level in prostatic tumor tissue and
 PT encoded polypeptides, useful for treating cancer and screening for
 PT therapeutic agents -
 XX Claim 1a; 107; 166pp; German.
 XX
 XX This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in prostatic tumor tissue and encode gene

CC products or their fragments. The products of the invention have
 CC antitumor activity. Polypeptides (i) encoded by (A) are used: (i) for
 CC identifying agents for treatment of prostatic cancer and (ii) for
 CC therapy of prostate cancer, optionally where expressed by gene therapy
 CC methods. (A) is also used to isolate full-length genes (for gene therapy)
 CC and for recombinant production of (I), which can be used to raise
 CC specific antibodies. (A) are identified by assembly of ESTs (expressed
 CC sequence tags) before they are analyzed for expression pattern (tissue
 CC specificity). This approach eliminates many of the false results, as
 CC regards tissue specificity, associated with known methods that use
 CC single (usually short) ESTs. AAZ33423-233476 represent expressed
 CC sequence tags described in the method of the invention.

XX Sequence 906 BP; 176 A; 303 C; 263 G; 164 T; 0 other;

Query Match 58.6%; Score 17; DB 20; Length 906;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
 |||||
 Db 228 ctgcagccatggctagg 244

RESULT 6
 AAI89867
 ID AAI89867 standard; cDNA; 1452 BP.

XX
 AC AAI89867;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 9927.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI; 2001-514838/56.
 DR
 DR P-PSDB; AAO09936.

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

PS Claim 1; SEQ ID NO 9927; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1452 BP; 291 A; 460 C; 429 G; 272 T; 0 other;

Query Match 58.6%; Score 17; DB 22; Length 1452;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
 |||||
 Db 207 ctgcagccatggctagg 223

RESULT 7
 AAF16096
 ID AAF16096 standard; cDNA; 1466 BP.

XX
 AC AAF16096;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:531.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease; ss.

XX Homo sapiens.

XX WO200055174-A1.
 XX
 PD 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587513/55.
 DR
 DR P-PSDB; AAB56893.

XX Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -

XX Claim 1; Page 993; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX SQ Sequence 1466 BP; 318 A; 453 C; 426 G; 269 T; 0 other;

Query Match 58.6%; Score 17; DB 21; Length 1466;
 Best Local Similarity 100.0%; Pred. No. 6.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 17; Conservative 0;

QY 6 ctgcagccatgctagg 22
 |||||
 Db 211 ctgcagccatgctagg 227

RESULT 8
 AAZ24399
 ID AAZ24399 standard; cDNA; 1467 BP.
 AC AAZ24399;
 XX
 DT 14-FEB-2000 (first entry)
 XX
 DE Human bladder tumour cDNA library derived EST 11.
 KW Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
 KW treatment; gene therapy; EST; ss.
 XX
 OS Homo sapiens.
 PN DE19818619-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 21-APR-1998; 98DE-1018619.
 XX
 PR 21-APR-1998; 98DE-1018619.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;
 XX WPI; 1999-612028/53.
 XX
 PT New nucleic acid sequences expressed in bladder tumor tissue, and
 PT derived polypeptides, for treatment of bladder tumor and identification
 PT of therapeutic agents -
 XX
 PS Claim 3; Page 68-69; 132pp; German.
 XX
 CC This invention describes novel polypeptide fragments (I) and the
 CC polynucleotides (II) that encode them that are highly expressed in a
 CC human bladder tumour and which have cytostatic activity. (II) are used
 CC for recombinant expression of (I) and to isolate complete genes. (I) are
 CC used to identify agents suitable for treatment of bladder cancer, to
 CC directly treat this form of cancer (including expression from gene
 CC therapy vectors) or are used in a preparation for cancer treatment. (I)
 CC is also used for the generation of specific antibodies. (II) are
 CC identified by assembling ESTs (expressed sequence tags) from a
 CC particular tissue type before comparison of expression patterns. This
 CC allows a significantly longer fragment of the gene to be revealed, and
 CC therefore reduces the number of failures associated with the fact that
 CC ESTs from different libraries may represent different parts of the same
 CC unknown gene, distorting the estimated frequency of occurrence in a
 CC particular tissue. AAZ43260-243309 represent expressed sequence tag (EST)
 CC fragments isolated from a human bladder tumour cDNA library which encode
 CC the proteins represented in AAZ66143-Y66198.
 XX
 SQ Sequence 1467 BP; 307 A; 463 C; 432 G; 265 T; 0 other;

Query Match 58.6%; Score 17; DB 20; Length 1467;
 Best Local Similarity 100.0%; Pred. No. 6.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 17; Conservative 0;

QY 6 ctgcagccatgctagg 22
 |||||
 Db 229 ctgcagccatgctagg 245

RESULT 9
 AAA79626
 ID AAA79626 standard; cDNA; 473 BP.
 AC AAA79626;
 XX
 DT 27-NOV-2000 (first entry)
 XX
 DE Pinus radiata cell signalling involved polynucleotide SEQ ID NO:427.
 XX
 KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism; ss.
 XX
 OS Pinus radiata.
 PN WO200042171-A1.
 XX
 PD 20-JUL-2000.
 XX
 PF 11-JAN-2000; 2000WO-US00724.
 XX
 PR 12-JAN-1999; 99US-0228986.
 PR 01-NOV-1999; 99US-0162866.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Strabala TJ, Nieuwenhuizen NJ;
 XX WPI; 2000-476052/41.
 XX
 PT Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 PT to external signals -
 XX
 PS Claim 1; Page 192; 527pp; English.
 XX
 CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,
 CC differentiation, elongation and survival, resistance to disease and
 CC nutrient metabolism. Examples of modifications which can be produced are
 CC altered fruit ripening and senescence of leaves and flowers e.g. to
 CC delay senescence and prolong the life of cut flowers or enhance
 CC senescence of reproductive organs to engineer sterile plants. Other
 CC modifications can be used to delay senescence in selected cell types or
 CC organs providing fruit and vegetables which have a longer shelf life
 CC between harvest and consumption, or to decrease branching frequency in
 CC forest tree species giving long stretches of valuable knot-free clear
 CC wood which can be used in solid timber furniture and veneers.
 XX
 SQ Sequence 473 BP; 105 A; 96 C; 118 G; 154 T; 0 other;

Query Match 55.2%; Score 16; DB 21; Length 473;
 Best Local Similarity 100.0%; Pred. No. 24; Mismatches 0; Indels 0; Gaps 0;
 Matches 16; Conservative 0;

QY 3 tatctgcagccatggc 18
 |||||
 Db 163 tatctgcagccatggc 178

```

RESULT 10
AAA79736
ID AAA79736 standard; cDNA; 4356 BP.
XX AC
XX AAA79736;
XX
XX 27-NOV-2000 (first entry)
XX
XX Pinus radiata cell signalling involved polynucleotide SEQ ID NO:888.
XX DE
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
XX KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
XX KW environmental change; development; cell proliferation; differentiation;
XX KW elongation; survival; disease resistance; nutrient metabolism; ss.
XX OS
XX Pinus radiata.
XX
XX WO200042171-A1.
XX PN
XX 20-JUL-2000.
XX PD
XX
XX 11-JAN-2000; 2000WO-US00724.
XX PF
XX 12-JAN-1999; 99US-0228986.
XX PR
XX 01-NOV-1999; 99US-0162866.
XX PR
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX PA
XX Strabala TJ, Nieuwenhuizen NJ;
XX PI
XX
XX WPI; 2000-476052/41.
XX DR
XX
XX Isolated polynucleotide encoding a polypeptide involved in cell
XX PT signalling used for generating transgenic plants with modified responses
XX PT to external signals -
XX
XX Claim 1; Page 440-442; 527pp; English.
XX PS
XX
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
XX CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
XX CC pine (Pinus radiata also known as Monterey pine). The protein sequences
XX CC are involved in cell signalling. The polynucleotide and protein
XX CC sequences can be used to modify the response of plant cells to external
XX CC signals e.g. environmental changes or pathogens during the growth and
XX CC development of a plant. They can be used to modify cell proliferation,
XX CC differentiation, elongation and survival, resistance to disease and
XX CC nutrient metabolism. Examples of modifications which can be produced are
XX CC altered fruit ripening and senescence of leaves and flowers e.g. to
XX CC delay senescence and prolong the life of cut flowers or enhance
XX CC senescence of reproductive organs to engineer sterile plants. Other
XX CC modifications can be used to delay senescence in selected cell types or
XX CC organs providing fruit and vegetables which have a longer shelf life
XX CC between harvest and consumption, or to decrease branching frequency in
XX CC forest tree species giving long stretches of valuable knot-free clear
XX CC wood which can be used in solid timber furniture and veneers.
XX
XX Sequence 4356 BP; 1273 A; 824 C; 901 G; 1358 T; 0 other;
XX SQ

```

```

Query Match 55.2%; Score 16; DB 21; Length 4356;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 tatctgcagccatggc 18
| | | | | | | | | |
Db 135 tatctgcagccatggc 150

```

```

RESULT 11
AAN90958/c
ID AAN90958 standard; DNA; 4481 BP.
XX AC
XX AAN90958;

```

```

XX 31-JAN-1990 (first entry)
XX DE
XX Sequence encoding germ cell alkaline phosphatase (ALP).
XX KW Germ cell alkaline phosphatase gene; placental ALP; cancer treatment;
XX OS
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX TATA_signal 368..374
XX FT /*tag= a
XX exon 450..516
XX FT /*tag= b
XX exon 611..727
XX FT /*tag= c
XX exon 839..954
XX FT /*tag= d
XX exon 1147..1321
XX FT /*tag= e
XX exon 1399..1565
XX FT /*tag= f
XX exon 1813..1947
XX FT /*tag= g
XX exon 2047..2119
XX FT /*tag= h
XX exon 2248..2382
XX FT /*tag= i
XX exon 2464..2655
XX FT /*tag= j
XX exon 2875..2991
XX FT /*tag= k
XX exon 3114..3409
XX FT /*tag= l
XX
XX WO8908254-A.
XX PN
XX 08-SEP-1989.
XX PD
XX
XX 08-FEB-1989; 89WO-US00567.
XX PF
XX 04-MAR-1988; 88US-0161138.
XX PR
XX (JOLL-) LA JOLLA CANCER RESEARCH FOUNDATION.
XX PA
XX Milan JL;
XX PI
XX WPI; 1989-278386/38.
XX DR P-PSDB; AAP01776.
XX DR
XX Germ cell alkaline phosphatase gene, protein and antibodies - used in
XX PT diagnosis and therapeutic treatment of certain cancers.
XX PT
XX Disclosure; fig 2; 31pp; English.
XX PS
XX
XX Sequence contains 11 exons (tag b - ) which when spliced together code
XX CC for germ cell alkaline phosphatase (ALP). The DNA was obt'd. from partial
XX CC MboI-digested human placental and splenic DNA genomic libraries
XX CC constructed in Charon 28 lambda phage. The DNA can be used to make probes
XX CC for diagnosis or antibodies to treat germ cell tumours where the ALP is
XX CC expressed. Germ cell ALP is also termed the "Nagao isoenzyme" and is a
XX CC placental-ALP (pLAP)-like isoenzyme showing 98% homology with PLAP.
XX
XX Sequence 4481 BP; 899 A; 1409 C; 1374 G; 799 T; 0 other;
XX SQ

```

```

Query Match 55.2%; Score 16; DB 10; Length 4481;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 gccatggcctaggccccc 26
| | | | | | | | | |
Db 1779 GCCATGGCTAGGCCCC 1764

```



```

RESULT 12
AAQ20514/c
ID AAQ20514 standard; DNA; 4487 BP.
XX AC AAQ20514;
XX DT 12-MAY-1992 (first entry)
XX DE Encodes germ cell alkaline phosphatase.
XX KW Placental ALP; cancer; ss.
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT GC_signal 148..156
XX FT GC_signal /*tag= a
XX FT GC_signal 315..323
XX FT /*tag= b
XX FT /*note= "exact match to SP1 binding site consensus"
XX FT GC_signal 339..347
XX FT /*tag= c
XX FT repeat_region 247..267
XX FT /*tag= d
XX FT repeat_unit 247..251
XX FT /*tag= e
XX FT /*rpt_type= DIRECT
XX FT TATA_signal 368..374
XX FT /*tag= f
XX FT exon 450..516
XX FT /*tag= g
XX FT intron /*number= 1
XX FT exon 517..610
XX FT /*tag= h
XX FT exon 611..727
XX FT /*tag= i
XX FT intron /*number= 2
XX FT exon 728..838
XX FT /*tag= j
XX FT exon 839..954
XX FT /*tag= k
XX FT intron /*number= 3
XX FT exon 955..1146
XX FT /*tag= l
XX FT exon 1147..1321
XX FT /*tag= m
XX FT intron /*number= 4
XX FT exon 1322..1398
XX FT /*tag= n
XX FT exon 1399..1571
XX FT /*tag= o
XX FT intron /*number= 5
XX FT exon 1572..1818
XX FT /*tag= p
XX FT exon 1819..1953
XX FT /*tag= q
XX FT intron /*number= 6
XX FT exon 1954..2052
XX FT /*tag= r
XX FT intron /*number= 7
XX FT exon 2053..2125
XX FT /*tag= s
XX FT intron 2126..2253
XX FT /*tag= t
XX FT exon 2254..2388
XX FT /*tag= u
XX FT intron /*number= 8
XX FT exon 2389..2469
XX FT /*tag= v
XX FT exon 2470..2661
XX FT /*tag= w
XX FT /*number= 9
XX FT 2662..2880
XX FT /*tag= x
XX FT 2881..2997
XX FT /*tag= y
XX FT /*number= 10
XX FT 2998..3119
XX FT /*tag= z
XX FT 3119..4487
XX FT /*tag= aa
XX FT /*number= 11
XX FT 1024..1030
XX FT /*tag= ab
XX FT GC_signal 1615..1622
XX FT /*tag= ac
XX FT GC_signal 1731..1739
XX FT /*tag= ad
XX FT GC_signal 1757..1765
XX FT /*tag= ae
XX FT GC_signal 1989..1996
XX FT /*tag= af
XX FT GC_signal complement (2205..2211)
XX FT /*tag= ag
XX FT GC_signal complement (2309..2316)
XX FT /*tag= ah
XX FT GC_signal 3054..3062
XX FT /*tag= ai
XX FT GC_signal complement (3555..3561)
XX FT /*tag= aj
XX FT GC_signal complement (3884..3890)
XX FT /*tag= ak
XX FT GC_signal complement (4187..4193)
XX FT /*tag= al
XX FT polyA_signal 4228..4233
XX FT /*tag= am
XX PN US5081227-A.
XX PD 14-JAN-1992.
XX XX
XX PF 04-MAR-1988; 88US-0164138.
XX PR 04-MAR-1988; 88US-0164138.
XX PA (LJOL-) LA JOLIA CANCER RES.
XX PI Millan JL;
XX DR WPI; 1992-048357/06.
XX DR P-PSDB; AAR20527.
XX PT New germ cell alkaline phosphatase and peptide fragments - giving
XX PT antibodies specific to germ cell alkaline phosphatase, used to
XX PT detect target, image germ cell and for site-directed therapy
XX PS Disclosure; Fig 2; 10pp; English.
XX CC Two different human genomic DNA libraries, one from placental tissue
XX CC and one from spleen tissue, were screened for the presence of
XX CC placental Alkaline Phosphatase (PALP)-related sequences. One clone
XX CC was present in both libraries and was chosen for further study. Its
XX CC sequence was compared to the known sequence of cDNA encoding PALP and
XX CC intron-exon boundaries identified by aligning the genomic sequence with
XX CC the cDNA. The new coding sequence is that of germ-line ALP.
XX CC The sequences listed in the Features Table as "GC-signals", although
XX CC highlighted in the specification, are not described in any detail
XX CC so there significance/function cannot be given.
XX SQ Sequence 4487 BP; 900 A; 1411 C; 1377 G; 799 T; 0 other;
Query Match 55.2%; Score 16; DB 13; Length 4487;

```

```

Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 gccatggttagggccc 26
DB 1785 GCCATGGCTAGGCCCC 1770

RESULT 13
ID AAL31467/c
AC AAL31467;
XX
XX
XX 24-JAN-2002 (first entry)
XX
XX Human SNP oligonucleotide #4675.
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX Homo sapiens.
XX
XX WQ200147944-A2.
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US35498.
XX
XX 28-DEC-1999; 99US-0173419.
XX
XX 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimketa RA, Leach M;
XX
XX WPI; 2001-465210/50.
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
XX Claim 1; Page 2731; 4143pp; English.
XX
XX The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX Sequence 51 BP; 14 A; 16 C; 12 G; 9 T; 0 other;

Query Match 51.7%; Score 15; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggct 19
DB 118 TCTGCAGCCATGGCT 104

us-09-810-861b-1.oligo.rng

Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggct 19
DB 47 TCTGCAGCCATGGCT 33

RESULT 14
ID AAX10671/c
AC AAX10671;
XX
XX 30-MAR-1999 (first entry)
XX
XX Human biallelic polymorphic DNA fragment WI-6303.
XX
XX Polymorphism: biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
XX
XX Homo sapiens.
XX
XX WQ9820165-A2.
XX
XX 14-MAY-1998.
XX
XX 05-NOV-1997; 97WO-US20313.
XX
XX 06-NOV-1996; 96US-0030455.
XX
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Hudson T, Lander ES, Wang D;
XX
XX WPI; 1998-286974/25.
XX
XX New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
XX
XX Claim 1; Page 70; 310pp; English.
XX
XX AAX10269-X12937 are human DNA fragments which contain biallelic
CC polymorphic markers which have been isolated using the primers
CC represented in AAX09121-X10268. The base occupying the polymorphic site
CC is indicated by the appropriate IUPAC-TUB ambiguity code. These fragments
CC can be used in methods for determining polymorphic forms in an individual
CC for use in e.g. forensics, paternity testing or for phenotypic typing for
CC diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC familial hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness; obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
XX
XX Sequence 204 BP; 47 A; 60 C; 43 G; 53 T; 1 other;

Query Match 51.7%; Score 15; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggct 19
DB 118 TCTGCAGCCATGGCT 104

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:26:31 ; Search time 6064.54 Seconds
(without alignments)
64.541 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatatctgagccatgctagggccccgc 29

Scoring table: OIRGO_NUC
Gap 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_esti.*
10: gb_est2.*
11: gb_htc.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pin.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Match	Length	DB	ID
1	17	58.6	305	9	AW404201
2	17	58.6	311	9	AA380508
3	17	58.6	323	10	BF803833
4	17	58.6	393	10	BE839073
5	17	58.6	422	9	AA401866
6	17	58.6	425	9	AW238889
7	17	58.6	440	10	BG140096
8	17	58.6	440	10	BE251243
9	17	58.6	441	9	AA314757
10	17	58.6	448	10	BF187245
11	17	58.6	456	10	BE313471
12	17	58.6	496	10	BE254863
13	17	58.6	501	10	BH021984
14	17	58.6	508	10	BF796488
15	17	58.6	510	9	AW247247
16	17	58.6	530	10	BE252091
17	17	58.6	534	10	BF309847

c	18	17	58.6	539	12	BH190610	BH190610 TC3-49C21
	19	17	58.6	551	10	BM021473	BM021473 1e77g09.y
	20	17	58.6	553	10	BE279799	BE279799 601157110
	21	17	58.6	565	10	BE395544	BE395544 601309888
	22	17	58.6	569	9	AA990251	AA990251 ua62b08.r
	23	17	58.6	570	10	BM010014	BM010014 603830775
	24	17	58.6	574	9	AW248246	AW248246 2819868.5
	25	17	58.6	574	10	BE259913	BE259913 601153452
	26	17	58.6	577	10	BE256672	BE256672 601107754
	27	17	58.6	580	10	BE252950	BE252950 601117541
	28	17	58.6	600	10	BI908883	BI908883 603066474
	29	17	58.6	604	10	BE396866	BE396866 601289671
	30	17	58.6	605	10	BE883652	BE883652 601507332
	31	17	58.6	607	10	BG334821	BG334821 602460814
	32	17	58.6	610	10	BE266003	BE266003 601194716
	33	17	58.6	616	10	BE267403	BE267403 601189637
	34	17	58.6	617	10	BG890568	BG890568 EST516419
	35	17	58.6	618	10	BI179157	BI179157 EST520102
	36	17	58.6	618	10	BE280185	BE280185 601158395
	37	17	58.6	619	10	BE251761	BE251761 60112458
	38	17	58.6	625	10	BE261435	BE261435 601147194
	39	17	58.6	634	10	BF689428	BF689428 602186787
	40	17	58.6	635	10	BE382514	BE382514 601297929
	41	17	58.6	638	10	BF312637	BF312637 601898128
	42	17	58.6	639	10	BE264391	BE264391 601191478
	43	17	58.6	643	10	BE617089	BE617089 601441620
	44	17	58.6	644	10	BE900214	BE900214 601673151
	45	17	58.6	646	10	BE885936	BE885936 601507567

ALIGNMENTS

RESULT 1

AW404201

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW404201 305 bp mRNA linear EST 16-FEB-2000
UI-HF-BLO-abq-b-08-0-UI.r1 NIH_MGC_37 Homo sapiens cDNA clone
IMAGE:3057351 5', mRNA sequence.

AW404201
GI:6923258

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 305)

NIH-MGC <http://mgi.nci.nih.gov/>

National Institutes of Health

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward

Location/Qualifiers

1..305

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3057351"

/clone_lib="NIH_MGC_37"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"

/note="Vector: pT7m3-Pac; Site_1: NotI; Site_2: Eco RI;

Constructed from size fractionated cytoplasmic mRNA

(1.5-2.5Kb). Directionally cloned. Cells provided by Louis

BASE COUNT 51 a 99 c 90 g 65 t
 ORIGIN M. Staudt, Ph.D. Library preparation by Maria de Fatima
 Ronaldo, Ph.D. and M. Bento Soares, Ph.D.*

Query Match 58.6%; Score 17; DB 9; Length 305;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatgctagg 22
 |||||
 DB 105 CTGACGCATGCTAGG 121

RESULT 2
 AA380508 311 bp mRNA linear EST 21-APR-1997
 LOCUS EST93479 Supt cells Homo sapiens cDNA 5' end, mRNA sequence.
 DEFINITION AA380508
 ACCESSION
 VERSION AA380508.1 GI:2032826
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 311)
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palao,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 JOURNAL based upon 83 million nucleotides of cDNA sequence
 MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
 COMMENT 96026280
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavetigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source
 1. .311
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="atcc:(inhost):184901"
 /db_xref="taxon:9606"
 /clone_lib="Supt cells"
 /cell_type="T-lymphocyte"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 58 a 93 c 87 g 71 t 2 others
 ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 311;

Best Local Similarity 100.0%; Pred. No. 60;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 ctgcagccatgctagg 22
 |||||
 DB 175 CTGACGCATGCTAGG 191

RESULT 3
 BF803833 323 bp mRNA linear EST 12-JAN-2001
 LOCUS MR0-CI0075-021100-003-a09 CI0075 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF803833
 ACCESSION BF803833
 VERSION BF803833.1 GI:12132822
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 323)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-CI0075-021100-003-a09&t3=2000-11-02&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 25
 High quality sequence stop: 323.

FEATURES
 source
 1. .323
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CI0075"
 /dev_stage="Adult"
 /note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 73 a 98 c 104 g 48 t
 ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 323;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatgctagg 22
 |||||
 DB 116 CTGACGCATGCTAGG 100

RESULT 4
 BE839073

```

LOCUS      BE839073      393 bp      mRNA      linear      EST 22-SEP-2000
DEFINITION RCO-FN0140-100700-022-a08 FN0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE839073
VERSION     BE839073.1 GI:10271555
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 393)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?li=st2-RCO-FN0140-100
            700-022-a08&t3=2000-07-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 393.
FEATURES    Location/Qualifiers
             1..393
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="FN0140"
             /dev_stage="Adult"
             /note="Organ: prostate; normal; Vector: puc18; Site_1: SmaI
             products derived from ORESTES PCR (U.S. Letters Patent
             application No. 196,716 - Ludwig Institute for Cancer
             Research) profiles into the puc 18 vector. Reverse
             transcription of tissue mRNA and cDNA amplification were
             performed under low stringency conditions."
BASE COUNT  76 a 129 c 108 g 80 t
ORIGIN
Query Match      58.6%; Score 17; DB 10; Length 393;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 131 CTGCAGCCATGGCTAGG 147

RESULT 5
AA401866
LOCUS      zv65c11.r1 Soares_total_fetus.Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:758516 5', mRNA sequence.
ACCESSION  AA401866
VERSION     AA401866.1 GI:2055885
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 422)
AUTHORS     Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
            Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            T., Waterston,K. and Wilson,R.
TITLE       WashU-Merck EST Project 1997
JOURNAL     Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL : contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 407.
FEATURES    Location/Qualifiers
             1..422
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="IMAGE:758516"
             /clone_lib="Soares_total_fetus.Nb2HF8_9w"
             /dev_stage="8-9 weeks"
             /lab_host="DH10B"
             /note="Vector: pT73D-Pac (Pharmacia) with a modified
             polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
             was prepared from mRNA obtained from pooled 8-9 week
             (total) fetus material with a Not I - oligo(NT) primer [5'
             TGTTACCAATCTGAGTGGGCGGCCGCTTAATTTTCTTTTCTTTT 3'].
             Double-stranded cDNA was ligated to Eco RI adaptors
             (Pharmacia), digested with Not I and cloned into the Not I
             and Eco RI sites of the modified pT73 vector. Library
             went through one round of normalization, and was
             constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT  80 a 129 c 123 g 90 t
ORIGIN
Query Match      58.6%; Score 17; DB 9; Length 422;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 174 CTGCAGCCATGGCTAGG 190

RESULT 6
AW238889
LOCUS      xb29c10.y1 NCI_CGAP_Lu31 Homo sapiens cDNA clone IMAGE:257714 5',
DEFINITION mRNA sequence.
ACCESSION  AW238889
VERSION     AW238889.1 GI:6571211
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 425)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL     Other_ESTs: xb29c10.x1
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: ATCC cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I. M. A. G. E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40RP from Gibco
 High quality sequence stop: 425.

FEATURES

SOURCE

1. .425
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2577714"
 /clone_lib="NCI_CGAP_Lu31"
 /sex="male"
 /dev_stage="fetal, 14 wk post-conception"
 /lab_host="DH10B"
 /note="Organ: lung, cell line; Vector: pCMV-SPORT6;
 Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally, no
 5' adaptor. Primer: Oligo dt. Full-length library
 constructed by Life Technologies." 87 t

BASE COUNT

76 a 143 c 119 g

ORIGIN

Query Match

Best Local Similarity 58.6%; Score 17; DB 9; Length 425;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22

|||||

Db 134 CTCGAGCCATGGCTAGG 150

RESULT 7

BG140096

LOCUS

DEFINITION EST480538 wild tomato pollen Lycopersicon pennellii cDNA clone EST 31-JAN-2001
 CLP16M11 5' sequence, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Lycopersicon pennellii.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE

AUTHORS

van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
 Hansen,C., Ronning,C. and Tanksley,S.

TITLE

JOURNAL

COMMENT

Unpublished (2001)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES

SOURCE

1. .440
 /organism="Lycopersicon pennellii"
 /cultivar="TA56"
 /db_xref="taxon:28526"
 /clone="cLPP16M11"
 /clone_lib="wild tomato pollen"
 /tissue_type="pollen"
 /dev_stage="pollen collected from open flowers"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; Pollen was collected from open flowers from
 L.pennellii TA56, and stored at -80 C until library
 construction." 90 t

BASE COUNT

118 a 170 c 61 g

ORIGIN

Query Match

58.6%; Score 17; DB 10; Length 440;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 atctgcagccatggcta 20

|||||

Db 13 ATCTGCAGCCATGGCTA 29

RESULT 8

BE251243

LOCUS

DEFINITION BE251243 440 bp mRNA linear EST 13-JUL-2000
 601116352F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3356964 5',
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 440)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLC1M63 row: g column: 13
 High quality sequence stop: 440.

Location/Qualifiers

1. .440

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3356964"

/clone_lib="NIH_MGC_16"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5',
 adaptor: GGCACGAG(C). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library." 88 t

FEATURES

SOURCE

80 a 147 c 125 g

Query Match 58.6%; Score 17; DB 10; Length 440;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22

|||||

Db 137 CTCGAGCCATGGCTAGG 153

RESULT 9

AA314757

LOCUS

DEFINITION AA314757 441 bp mRNA linear EST 19-APR-1997
 EST186380 Colon carcinoma (HCC) cell line II Homo sapiens cDNA 5',
 end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 human.


```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 441)
REFERENCE
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Whai,C., Clayton,R.A.,
Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
Kelley,J.C., Liu,L.-T., Marmaros,S.M., Merrick,J.M.,
Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudet,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
JOURNAL based upon 83 million nucleotides of cDNA sequence
MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
COMMENT 96026280
Other ESTs: THC166913
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES
source Location/Qualifiers
1..441
/organism="Homo sapiens"
/db_xref="ATCC (inhost):111318"
/db_xref="taxon:9606"
/clone_lib="Colon carcinoma (HCC) cell line II"
/tissue_type="colon"
/cell_type="KM12C"
/contig="KM12C"
/contig_line="KM12C(HCC)-parental human colon carcinoma
;Dukes B2"
/notes="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 84 a 143 c 124 g 89 t 1 others
ORIGIN
Query Match 58.6%; Score 17; DB 9; Length 441;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ctgcagccatggctagg 22
|||||
Db 184 CTGCAGCCATGGCTAGG 200
RESULT 10
BF187245 448 bp mRNA linear EST 01-NOV-2000
LOCUS EST443532 potato stolon, Cornell University Solanum tuberosum cDNA
DEFINITION clone cSTA39J13 5' sequence, mRNA sequence.
ACCESSION BF187245
VERSION BF187245.1 GI:11069464
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

```

```

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 448)
REFERENCE
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: TIGR
The Institute for Genomic Research
Rockville, MD 20850, USA
Tel: (301)838-0200
Email: http://www.tigr.org/tdb/potato/
For clone information, please contact Research Genetics, Libraries
Division, tel# 1-800-711-6195, email= cdna@resgen.com.
FEATURES
source Location/Qualifiers
1..448
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone_lib="cSTA39J13"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early
stages of tuber formation."
BASE COUNT 120 a 178 c 62 g 88 t
ORIGIN
Query Match 58.6%; Score 17; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 atctgcagccatggcta 20
|||||
Db 13 ATCTGCAGCCATGGCTA 29
RESULT 11
BE313471 456 bp mRNA linear EST 26-OCT-2000
LOCUS 601148429F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163854 5',
DEFINITION mRNA sequence.
ACCESSION BE313471
VERSION BE313471.1 GI:9133413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 456)
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

```

Plate: LICM125 row: 1 column: 07
High quality sequence stop: 455.

FEATURES

source

1. .456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3163854"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
82 a 154 c 128 g 92 t

BASE COUNT
ORIGIN

Query Match 58.68; Score 17; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatgctagg 22

|||||

Db 144 CTGCAGCCATGCTAGG 160

RESULT 12

BE254863

LOCUS

DEFINITION 60111187f1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3352124 5', linear EST 13-JUL-2000

ACCSSION BE254863

VERSION BE254863.1 GI:9125296

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 496)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgep@remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LICM150 row: m column: 21

High quality sequence stop: 496.

FEATURES

source

1. .496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3352124"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
90 a 174 c 139 g 98 t

BASE COUNT

ORIGIN

BASE COUNT
ORIGIN

91 a 172 c 138 g 95 t

Query Match

58.68; Score 17; DB 10; Length 496;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatgctagg 22

|||||

Db 105 CTGCAGCCATGCTAGG 121

RESULT 13

BM021984

LOCUS

DEFINITION IE74f02.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens cDNA 5', mRNA sequence.

ACCSSION BM021984

VERSION BM021984.1 GI:16536340

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 501)

AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas

, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T.,

Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brown@fas.harvard.edu)

High quality sequence stop: 434.

Location/Qualifiers

1. .501

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPOT1; Site_1: Not 1;

Site_2: Sal 1; Starting library constructed using

Superscript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size selected by column

fractionation; average insert size 1.08 kb. Library was

amplified once on solid support and plasmid DNA from

library was prepared. The library DNA was normalized by

method #4 from Bonaldo, Lennon, and Soares 1996 Genome

Research 6:791-806; 0.5 microgram single-stranded library

plasmid DNA was mixed with 5 micrograms PCR product

representing library inserts and hybridized to an EcoT of

20. Single-stranded (unhybridized) plasmids were isolated

by hydroxyapatite chromatography and used to make this

library."

90 a 174 c 139 g 98 t

BASE COUNT

ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 501;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
 |||||
 Db 131 CTGCAGCCATGGCTAGG 147

RESULT 14
 BF796488
 LOCUS 508 bp mRNA linear EST 12-JAN-2001
 DEFINITION 602259947F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4343243 5',
 mRNA sequence.
 ACCESSION BF796488
 VERSION BF796488.1 GI:12101542
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 508)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9959 row: n column: 12
 High quality sequence stop: 505.

FEATURES
 source
 1..508
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4343243"
 /clone_lib="NIH_MGC_85"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 92 a 172 c 143 g 101 t
 ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 508;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
 |||||
 Db 160 CTGCAGCCATGGCTAGG 176

RESULT 15.
 AW247247
 LOCUS 510 bp mRNA linear EST 07-JAN-2000
 DEFINITION 2820405.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820405 5',
 mRNA sequence.
 ACCESSION AW247247
 VERSION AW247247.1 GI:6590240
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Other_ESTs: 2820405.3prime
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: DCTD/DRP CDNA Library Preparation: Ling
 Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
 Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
 project Clone distribution: MGC clone distribution information can
 be found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
 Scores: PHRED from University of Washington Genome Center. Vector
 Trimming: cross_match from University of Washington Genome Center
 PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
 Drosophila Genome Project. University of Washington Genome Center:
 http://www.genome.washington.edu
 Plate: LLCM4 row: B column: 22
 High quality sequence stop: 458.

FEATURES
 source
 1..510
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2820405"
 /clone_lib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5',
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 92 a 167 c 150 g 101 t
 ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 510;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
 |||||
 Db 180 CTGCAGCCATGGCTAGG 196

Search completed: August 31, 2002, 19:58:44
 Job time: 16333 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 17:15:36 ; Search time 154.73 Seconds
(without alignments)
46.037 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattcgagccatggctagggcccccgc 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14	48.3	22	1 US-08-104-073-13	Sequence 13, Appl
C 2	14	48.3	24	1 US-08-105-483-50	Sequence 50, Appl
C 3	14	48.3	24	1 US-08-105-483-51	Sequence 51, Appl
C 4	14	48.3	24	1 US-08-709-209-50	Sequence 50, Appl
C 5	14	48.3	24	1 US-08-709-209-51	Sequence 51, Appl
C 6	14	48.3	24	1 US-08-458-101-50	Sequence 50, Appl
C 7	14	48.3	24	1 US-08-458-101-51	Sequence 51, Appl
C 8	14	48.3	107	2 US-08-893-049-6	Sequence 6, Appl
C 9	14	48.3	292	2 US-08-230-928-2	Sequence 2, Appl
C 10	14	48.3	292	2 US-08-477-978-6	Sequence 6, Appl
C 11	14	48.3	1774	1 US-08-377-293-3	Sequence 3, Appl
C 12	14	48.3	2068	4 US-09-318-448-15	Sequence 15, Appl
C 13	14	48.3	2188	1 US-07-706-872-2	Sequence 2, Appl
C 14	14	48.3	2505	1 US-08-391-615-1	Sequence 1, Appl
C 15	13	44.8	18	4 US-08-718-388-18	Sequence 18, Appl
C 16	13	44.8	19	4 US-08-718-388-22	Sequence 22, Appl
C 17	13	44.8	28	3 US-09-009-156-12	Sequence 12, Appl
C 18	13	44.8	28	4 US-09-372-154-12	Sequence 12, Appl
C 19	13	44.8	32	3 US-08-722-719-25	Sequence 25, Appl
C 20	13	44.8	42	3 US-08-836-337-1	Sequence 1, Appl
C 21	13	44.8	194	2 US-08-634-797-48	Sequence 48, Appl
C 22	13	44.8	239	4 US-09-328-869-15	Sequence 15, Appl
C 23	13	44.8	239	4 US-09-629-774A-15	Sequence 15, Appl
C 24	13	44.8	244	4 US-09-034-205-29	Sequence 29, Appl
C 25	13	44.8	244	4 US-08-934-097A-29	Sequence 29, Appl
C 26	13	44.8	244	4 US-08-851-588-29	Sequence 29, Appl
C 27	13	44.8	244	4 US-09-677-218B-29	Sequence 29, Appl

C 28	13	44.8	244	4 US-09-677-192-29	Sequence 29, Appl
C 29	13	44.8	281	2 US-08-757-653-126	Sequence 126, App
C 30	13	44.8	281	2 US-08-757-653-132	Sequence 132, App
C 31	13	44.8	289	4 US-09-034-205-23	Sequence 23, Appl
C 32	13	44.8	289	4 US-08-934-097A-23	Sequence 23, Appl
C 33	13	44.8	289	4 US-08-851-588-23	Sequence 23, Appl
C 34	13	44.8	289	4 US-09-677-218B-23	Sequence 23, Appl
C 35	13	44.8	289	4 US-09-677-192-23	Sequence 23, Appl
C 36	13	44.8	691	2 US-08-365-486A-29	Sequence 29, Appl
C 37	13	44.8	691	4 US-08-880-342-29	Sequence 29, Appl
C 38	13	44.8	730	4 US-09-342-084-5	Sequence 5, Appl
C 39	13	44.8	803	1 US-08-157-235-1	Sequence 1, Appl
C 40	13	44.8	803	1 US-08-157-235-2	Sequence 2, Appl
C 41	13	44.8	803	1 US-08-157-235-3	Sequence 3, Appl
C 42	13	44.8	993	4 US-08-959-004-2	Sequence 2, Appl
C 43	13	44.8	1115	1 US-08-190-802A-19	Sequence 19, Appl
C 44	13	44.8	1115	4 US-08-477-346-19	Sequence 19, Appl
C 45	13	44.8	1115	4 US-08-473-089-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-104-073-13/c
; Sequence 13, Application US/08104073
; Patent No. 5589610
; GENERAL INFORMATION:
; APPLICANT: De Beuckeleer, Marc
; APPLICANT: Herdies, Lydia
; APPLICANT: Gossele, Veronique
; APPLICANT: Mariana, Celestina
; TITLE OF INVENTION: Stamen-Specific Promoters from Corn
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5589610west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,073
; FILING DATE: 05-AUG-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00275
; FILING DATE: 05-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91401787.6
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91400300.9
; FILING DATE: 07-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.92USWO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-104-073-13

Query Match 48.3%; Score 14; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggct 19
|||||
DB 22 CTCAGCCATGGCT 9

RESULT 2

US-08-105-483-50
; Sequence 50, Application US/08105483
; Patent No. 5494807

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford

; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue

; CITY: New York
; STATE: NY

; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,483

; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951

; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454310-2400
; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; US-08-105-483-50

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggc 18
|||||
DB 5 TCTGCAGCCATGGC 18

RESULT 3

US-08-105-483-51/c

; Sequence 51, Application US/08105483
; Patent No. 5494807

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo

; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
;

; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,483
; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-105-483-51

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggc 18
|||||
DB 24 TCTGCAGCCATGGC 11

RESULT 4

US-08-709-209-50

; Sequence 50, Application US/08709209
; Patent No. 5762938

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo

; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; NUMBER OF SEQUENCES: 462

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford

; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue

; CITY: New York
; STATE: NY

; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,209

; FILING DATE: 21-AUG-1996
; CLASSIFICATION: 424

;; PRIOR APPLICATION DATA: US 08/105,483
;; FILING DATE: 12-AUG-1993
;; APPLICATION NUMBER: US 07/847,951
;; FILING DATE: 06-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frommer, William S.
;; REGISTRATION NUMBER: 25,506
;; REFERENCE/DOCKET NUMBER: 454310-2400
;; TELEPHONE: (212) 840-3333
;; TELEFAX: (212) 840-0712
;; INFORMATION FOR SEQ ID NO: 50:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 24 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-709-209-50

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagccatggc 18
|||||
Db 5 TCTGCAGCCATGGC 18

RESULT 5
US-08-709-209-51/c
;; Sequence 51, Application US/08709209
;; Patent No. 5762938
;; GENERAL INFORMATION:
;; APPLICANT: Paoletti, Enzo
;; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
;; TITLE OF INVENTION: STRAIN
;; NUMBER OF SEQUENCES: 462
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Curtis, Morris & Safford
;; ADDRESSEE: c/o William S. Frommer
;; STREET: 530 Fifth Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/709,209
;; FILING DATE: 21-AUG-1996
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/105,483
;; FILING DATE: 12-AUG-1993
;; APPLICATION NUMBER: US 07/847,951
;; FILING DATE: 06-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frommer, William S.
;; REGISTRATION NUMBER: 25,506
;; REFERENCE/DOCKET NUMBER: 454310-2400
;; TELEPHONE: (212) 840-3333
;; TELEFAX: (212) 840-0712
;; INFORMATION FOR SEQ ID NO: 51:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 24 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single

;; TOPOLOGY: linear
US-08-709-209-51

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagccatggc 18
|||||
Db 24 TCTGCAGCCATGGC 11

RESULT 6
US-08-458-101-50
;; Sequence 50, Application US/08458101
;; Patent No. 5766599
;; GENERAL INFORMATION:
;; APPLICANT: Paoletti, Enzo
;; APPLICANT: Perkus, Marion E.
;; APPLICANT: Taylor, Jill
;; APPLICANT: Tartaglia, James
;; APPLICANT: No. 5766599ton, Elizabeth K.
;; APPLICANT: Riviere, Michel
;; APPLICANT: de Taisne, Charles
;; APPLICANT: Limbach, Keith J.
;; APPLICANT: Johnson, Gerard P.
;; APPLICANT: Pincus, Steven E.
;; APPLICANT: Cox, William I.
;; APPLICANT: Audonnet, Jean-Christophe Francis
;; APPLICANT: Gettig, Russell Robert
;; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
;; TITLE OF INVENTION: STRAIN
;; NUMBER OF SEQUENCES: 467
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Curtis, Morris & Safford
;; ADDRESSEE: c/o William S. Frommer
;; STREET: 530 Fifth Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/458,101
;; FILING DATE: 01-JUN-1995
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frommer, William S.
;; REGISTRATION NUMBER: 25,506
;; REFERENCE/DOCKET NUMBER: 454310-2740
;; TELEPHONE: (212) 840-3333
;; TELEFAX: (212) 840-0712
;; INFORMATION FOR SEQ ID NO: 50:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 24 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-458-101-50

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagccatggc 18
|||||

```

Db      5  TCTGCAGCCATGCC 18

RESULT 7
US-08-458-101-51/c
; Sequence 51, Application US/08458101
; Patent No. 5766599
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: Perkus, Marlon E.
; APPLICANT: Taylor, Jill
; APPLICANT: Tartaglia, James
; APPLICANT: No. 5766599ton, Elizabeth K.
; APPLICANT: Riviere, Michel
; APPLICANT: de Taisne, Charles
; APPLICANT: Limbach, Keith J.
; APPLICANT: Johnson, Gerard P.
; APPLICANT: Pincus, Steven E.
; APPLICANT: Cox, William I.
; APPLICANT: Audonnet, Jean-Christophe Francis
; APPLICANT: Gettig, Russell Robert
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,101
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-458-101-51

Query Match      48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5  tctgcagccatgcc 18
        |||||
Db      24  TCTGCAGCCATGCC 11

RESULT 8
US-08-893-049-6
; Sequence 6, Application US/08893049
; Patent No. 5962769
; GENERAL INFORMATION:
; APPLICANT: HOWARD, John A.
; APPLICANT: ALBERTSEN, Marc C.
;

; TITLE OF INVENTION: INDUCTION OF MALE STERILITY IN PLANTS BY
; TITLE OF INVENTION: EXPRESSION OF HIGH LEVELS OF AVIDIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,049
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,582
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/314/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-893-049-6

Query Match      48.3%; Score 14; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3  tatctgcagccatg 16
        |||||
Db      94  TATCTGCAGCCATG 107

RESULT 9
US-08-230-929-2
; Sequence 2, Application US/08230929
; Patent No. 5470359
; GENERAL INFORMATION:
; APPLICANT: Huffman, Gary A.
; TITLE OF INVENTION: REGULATORY ELEMENT CONFERRING TAPETUM
; TITLE OF INVENTION: SPECIFICITY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,929
; FILING DATE: 21-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.

```


REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/210/PIHI
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-230-929-2

Query Match 48.3%; Score 14; DB 1; Length 292;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatg 16
|||||

Db 279 TATCTGCAGCCATG 292

RESULT 10
US-08-477-978-6
; Sequence 6, Application US/08477978
; Patent No. 5837850
; GENERAL INFORMATION:
; APPLICANT: HUFFMAN, Gary A.
; TITLE OF INVENTION: REGULATORY ELEMENT CONFERRING TAPETUM
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,978
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/230,929
; FILING DATE: 21-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/330/PIHI
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-477-978-6

Query Match 48.3%; Score 14; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatg 16
|||||

Db 279 TATCTGCAGCCATG 292

RESULT 11
US-08-377-292-3/c
; Sequence 3, Application US/08377292
; Patent No. 5693615
; GENERAL INFORMATION:
; APPLICANT: STONE, ROGER L.
; TITLE OF INVENTION: THERAPEUTIC FORMULAS FOR OSTEOINDUCTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 45239-8707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,292
; FILING DATE: 23-JAN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,435
; FILING DATE:
; APPLICATION NUMBER: US/08/117,367
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corstanje, Brahm J.
; REGISTRATION NUMBER: 34,804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-245-2858
; TELEFAX: 513-741-3012
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-377-292-3

Query Match 48.3%; Score 14; DB 1; Length 1774;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 cagccatggctagg 22
|||||

Db 327 CAGCCATGGCTAGG 314

RESULT 12
US-09-318-448-15
; Sequence 15, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2068
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-15

Query Match 48.3%; Score 14; DB 4; Length 2068;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggct 19
Db 68 ctgcagccatggct 81
|||||

RESULT 13

US-07-706-872-2
; Sequence 2, Application US/07706872
; Patent No. 5237056
; GENERAL INFORMATION:
; APPLICANT: Fischbach, Gerald D.
; TITLE OF INVENTION: Identification of a Protein Which
; TITLE OF INVENTION: Promotes the Synthesis of Acetylcholine Receptors and Uses
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07706,872
; FILING DATE: 19910529
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU91-01
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 2188 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 172..972
US-07-706-872-2

Query Match 48.3%; Score 14; DB 1; Length 2188;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 agccatggctagcc 23
Db 168 AGCCATGGCTAGGC 181
|||||

RESULT 14

US-08-391-615-1/c
; Sequence 1, Application US/08391615
; Patent No. 5550054
; GENERAL INFORMATION:
; APPLICANT: Witte, Owen
; APPLICANT: Tsukada, Satoshi
; APPLICANT: Saffran, Douglas
; APPLICANT: Rawlings, David
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,615
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/006,449
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 137..2116
US-08-391-615-1

Query Match 48.3%; Score 14; DB 1; Length 2505;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggct 19
Db 146 CTGCAGCCATGGCT 133
|||||

RESULT 15

US-08-718-388-18
; Sequence 18, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:
; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING IgG Fc REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747

```

; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,388
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "oligonucleotide"
US-08-718-388-18

```

```

Query Match 44.88; Score 13; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 tctgcagccatgg 17
Db 5 TCTGCAGCCATGC 17

```

Search completed: August 31, 2002, 22:37:20
Job time: 19304 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:34:43 ; Search time 9324.54 Seconds
(without alignments)
69.572 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggctactatcaggtagcgtgagcaatttg 31

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :		GenEmbl:	
1:	gb_ba:*	31	100.0
2:	gb_hlg:*	31	100.0
3:	gb_in:*	31	100.0
4:	gb_om:*	18	58.1
5:	gb_ov:*	18	58.1
6:	gb_pat:*	18	58.1
7:	gb_ph:*	18	58.1
8:	gb_pl:*	18	58.1
9:	gb_pr:*	18	58.1
10:	gb_ro:*	18	58.1
11:	gb_sts:*	18	58.1
12:	gb_sv:*	18	58.1
13:	gb_un:*	18	58.1
14:	gb_vl:*	18	58.1
15:	em_ba:*	17	54.8
16:	em_fun:*	17	54.8
17:	em_hum:*	17	54.8
18:	em_in:*	17	54.8
19:	em_mu:*	17	54.8
20:	em_om:*	17	54.8
21:	em_ot:*	17	54.8
22:	em_ov:*	17	54.8
23:	em_pat:*	17	54.8
24:	em_ph:*	17	54.8
25:	em_pl:*	17	54.8
26:	em_ro:*	17	54.8
27:	em_sts:*	17	54.8
28:	em_un:*	17	54.8
29:	em_vl:*	17	54.8
30:	em_hlg_hum:*	17	54.8
31:	em_hlg_inv:*	17	54.8
32:	em_hlg_other:*	17	54.8
33:	em_htgo_inv:*	17	54.8

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result	Query		
No.	Score	Match Length DB ID	Description

ALIGNMENTS

RESULT 1		AX275253		31 bp		DNA		linear		PAT 29-OCT-2001	
AX275253		Sequence 2 from Patent WO0171014.									
LOCUS		AX275253									
DEFINITION		AX275253									
ACCESSION		AX275253									
VERSION		AX275253.1									
KEYWORDS		synthetic construct.									
SOURCE		synthetic construct.									
ORGANISM		artificial sequence.									
REFERENCE		1 (sites)									
AUTHORS		Mor.T., Soreq.H., Arntzen,C. and Mason,H.									
TITLE		Expression of recombinant human acetylcholinesterase in transgenic plants									
JOURNAL		Patent: WO 0171014-A 2 27-SEP-2001;									
		BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor.									
		Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,									
		Hugh S. (US)									
FEATURES		Location/Qualifiers									
		1..31									
		/organism="synthetic construct"									
		/db_xref="taxon:3630"									
		/note="primer pAChE-kpn, derived from human AChE gene and									

AX275253	Sequence	AX275253	31	100.0	31	6	AX275253
AX275254	Sequence	AX275254	6	5767	6	6	AX275254
AX275255	Sequence	AX275255	6	14446	6	6	AX275255
AC005975	Drosophila	AC005975	2	23223	2	2	AC005975
AC013153	Drosophila	AC013153	2	146775	2	2	AC013153
AC099035	Drosophila	AC099035	3	175451	3	3	AC099035
AC008352	Drosophila	AC008352	3	181463	3	3	AC008352
AE003459	Drosophila	AE003459	3	299751	3	3	AE003459
AF099920	Caenorhab	AF099920	3	34690	3	3	AF099920
AC044838	Mus muscu	AC044838	2	178881	2	2	AC044838
AL606531	Mus muscu	AL606531	2	205165	2	2	AL606531
AC006719	Caenorhab	AC006719	2	209365	2	2	AC006719
AL645602	Mus muscu	AL645602	2	271254	2	2	AL645602
U34962	Human trans	U34962	9	1019	9	9	U34962
AB027318	Cordyceps	AB027318	8	2235	8	8	AB027318
AB027322	Cordyceps	AB027322	8	2958	8	8	AB027322
U23511	Xanthomonas	U23511	1	16075	1	1	U23511
AC084295	Oryza sat	AC084295	8	89172	8	8	AC084295
AC023660	Homo sapi	AC023660	2	113392	2	2	AC023660
AC099323	Oryza sat	AC099323	2	130049	2	2	AC099323
AL513467	Neurospor	AL513467	8	NC17E5	8	8	AL513467
AL669924	Homo sapi	AL669924	2	136657	2	2	AL669924
AC008412	Homo sapi	AC008412	2	137545	2	2	AC008412
AC007623	Homo sapi	AC007623	2	163275	2	2	AC007623
AC091787	Oryza sat	AC091787	2	170518	2	2	AC091787
AC007834	Homo sapi	AC007834	2	213891	2	2	AC007834
AC102953	Homo sapi	AC102953	2	225506	2	2	AC102953
L22561	Human acety	L22561	9	510	9	9	HUMACHE03
M75539	Human acety	M75539	9	535	9	9	HUMACHEA
M76540	Mouse acety	M76540	10	840	10	10	MUSACHEA
Y16305	Trichoderma	Y16305	8	1024	8	8	TRQID74PR
U92014	Human clone	U92014	9	1415	9	9	HSU92014
AF278605	Homo sapi	AF278605	9	1641	9	9	AF278605
AX275256	Sequence	AX275256	6	1725	6	6	AX275256
AF007533	Coemansia	AF007533	8	1838	8	8	AF007533
AB027512	Lentinula	AB027512	8	2050	8	8	AB027512
E16408	Lentinus ed	E16408	6	2051	6	6	E16408
S50879	acetylcholin	S50879	10	2066	10	10	S50879
AX306139	Sequence	AX306139	6	2089	6	6	AX306139
X56518	Mouse mRNA	X56518	10	2089	10	10	WMACHE
M5040	Human acety	M5040	9	2218	9	9	HUMACHE
AR070205	Sequence	AR070205	6	2256	6	6	AR070205
AC002251	Homo sapi	AC002251	9	2281	9	9	AC002251

```

BASE COUNT      7 a      7 c      9 g      8 t
ORIGIN
modified to introduce a Kpn I restriction site"

Query Match      100.0%; Score 31; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtagctatcaggtagcgctgagcaatttg 31
|||||
Db 1 CGGTACCTATCAGGTAGCGCTGAGCAATTG 31

RESULT 2
AX275254/c
LOCUS      AX275254      5767 bp      DNA      linear      PAT 29-OCT-2001
DEFINITION Sequence 3 from Patent WO0171014.
ACCESSION  AX275254
VERSION     AX275254.1 GI:16547674
KEYWORDS   synthetic construct.
SOURCE     synthetic construct
ORGANISM   artificial sequence.

REFERENCE  1 (sites)
AUTHORS   Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE     Expression of recombinant human acetylcholinesterase in transgenic
JOURNAL   plants
PATENT    Patent: WO 0171014-A 3 27-SEP-2001;
          BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor.
          Tsaafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
          Hugh S. (US)
FEATURES   Location/Qualifiers
            source
              1..5767
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="plasmid vector pTW034"

BASE COUNT      1390 a      1521 c      1495 g      1361 t
ORIGIN

Query Match      100.0%; Score 31; DB 6; Length 5767;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtagctatcaggtagcgctgagcaatttg 31
|||||
Db 2571 CGGTACCTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
AX275255/c
LOCUS      AX275255      14446 bp      DNA      linear      PAT 29-OCT-2001
DEFINITION Sequence 4 from Patent WO0171014.
ACCESSION  AX275255
VERSION     AX275255.1 GI:16547675
KEYWORDS   synthetic construct.
SOURCE     synthetic construct
ORGANISM   artificial sequence.

REFERENCE  1 (sites)
AUTHORS   Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE     Expression of recombinant human acetylcholinesterase in transgenic
JOURNAL   plants
PATENT    Patent: WO 0171014-A 4 27-SEP-2001;
          BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor.
          Tsaafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
          Hugh S. (US)
FEATURES   Location/Qualifiers
            source
              1..14446
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="plasmid vector pTW036"

BASE COUNT      3231 a      3831 c      4046 g      3042 t      296 others
ORIGIN
modified to introduce a Kpn I restriction site"

Query Match      100.0%; Score 31; DB 6; Length 14446;
Best Local Similarity 100.0%; Pred. No. 5e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtagctatcaggtagcgctgagcaatttg 31
|||||
Db 1875 CGGTACCTATCAGGTAGCGCTGAGCAATTG 1845

RESULT 4
AC005975
LOCUS      AC005975      23223 bp      DNA      linear      HTG 30-JUL-1999
DEFINITION Drosophila melanogaster chromosome 2 clone DS02840 (D428) map
          59C3-59C5 strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***, 2
          unordered pieces.
ACCESSION  AC005975
VERSION     AC005975.3 GI:5656715
KEYWORDS   HTG; HTGS.PHASE1.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster

REFERENCE  1 (bases 1 to 23223)
AUTHORS   Celnikier,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
          Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
          Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.D.,
          Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
          Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
          Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
          Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
          Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
          Rubin,G.M.
TITLE     Sequencing of Drosophila melanogaster
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 23223)
AUTHORS   Celnikier,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
          Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
          Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.D.,
          Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
          Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
          Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
          Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
          Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
          Zieran,L.L. and Rubin,G.M.
TITLE     Direct Submission
JOURNAL   Submitted (19-NOV-1998) Drosophila Genome Center, Lawrence Berkeley
          Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT    On Jul 30, 1999 this sequence version replaced gi:5597039.
          For further information about this sequence, including its location
          and relationship to other sequences, please visit our sequence
          archive Web site (http://www.fruitfly.org/sequence/) or send email
          to bdgpe@fruitfly.berkeley.edu. All contigs in this submission meet
          the following cutoffs: length >= 200 bases. P1 library location:
          30-56.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 2 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          * 1 22689: contig of 22689 bp in length
          * 22690 22769: gap of unknown length
          * 22770 23223: contig of 454 bp in length.
          * Location/Qualifiers
            1..23223
            /organism="Drosophila melanogaster"

```

```

/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2"
/map="59c3-59c5"
/clone="DS02840 (D428)"
/clone.lib="p1 library, partial Sau3A in pAd10sacBII"

6308 a 5342 c 5216 g 6252 t 105 others

BASE COUNT
ORIGIN

Query Match      58.1%; Score 18; DB 2; Length 23223;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 gtacgcctgagcaattg 31
|||||
Db 2921 GTAGCGCTGAGCAATTG 2938

RESULT 5
AC013153
LOCUS
DEFINITION
Drosophila melanogaster, 146775 bp DNA linear HTG 03-NOV-1999
pieces.
ACCESSION AC013153
VERSION AC013153.1 GI:6223178
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 146775)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10213761 by the submitter.
For further information on this sequence you may e-mail to
fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1..146775
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 40744 a 32885 c 33046 g 40100 t
ORIGIN

Query Match      58.1%; Score 18; DB 2; Length 146775;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 gtacgcctgagcaattg 31
|||||
Db 57937 GTAGCGCTGAGCAATTG 57954

RESULT 6
AC099035
LOCUS
DEFINITION
Drosophila melanogaster, 175451 bp DNA linear INV 08-NOV-2001
BACR37H13, complete sequence.
ACCESSION AC099035
VERSION AC099035.1 GI:16798967
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

```

REFERENCE AUTHORS

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 175451)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferriera,S., Frise,E., Galle,R.F., Hoskins,R.A., Hostin,D., Howland,T.J.,
Gonzalez,M., Houck,J., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
Ibegwam,C., Jalali,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
McIntosh,T.C., Moy,M., Park,S., Patel,S., Pfeiffer,B.,
Pacleb,J., Paragas,V., Park,S., Pittman,G.S., Richards,S., Scheeler,F.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

TITLE JOURNAL REFERENCE AUTHORS

Unpublished
Sequencing of Drosophila chromosome 2R, region 59B-59C

2 (bases 1 to 175451)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferriera,S., Frise,E., Galle,R.F., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

TITLE JOURNAL

Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

COMMENT

This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgpf@fruitfly.berkeley.edu.

FEATURES source

1..175451
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="59B-59C"
/clone="BACR37H13 (DI380)"
/clone.lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBAC3.6)"
BASE COUNT 48153 a 39588 c 39612 g 48098 t
ORIGIN

Query Match 58.1%; Score 18; DB 3; Length 175451;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 gtacgcctgagcaattg 31

|||||
Db 21108 GTAGCGCTGAGCAATTG 21125

RESULT 7

AC008352/c
LOCUS AC008352 181463 bp DNA linear INV 27-APR-2001
DEFINITION Drosophila melanogaster, chromosome 2R, region 59C-59D, BAC clone

CDS	<p>/map="59B2-59B3"</p> <p>/db_xref="FLYBASE:FBan0003501"</p> <p>/db_xref="FLYBASE:FBgn0034791"</p> <p>complement(join(662..1075,1557..1754))</p> <p>/gene="CG3501"</p> <p>/note="CG3501 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003501"</p> <p>/db_xref="FLYBASE:FBgn0034791"</p> <p>/protein_id="AAF46921.1"</p> <p>/db_xref="GI:7291496"</p> <p>/translation="MCDYKVSERAYAKLIFHAQYPHQAVNGLLLAEKTSKGSOVEIV DAIPLHQCLYVTPMAEVALMLIDAHAEERGLVIAGYAAAPENYDNOVQKTPAAKIA DKIQENKACFFVVDKMLTLQHDRAAIQVFNCPGSGARWSKAKETLQASDTLEG VSLLUKKGAMRDLDFONHLDNDPKNTDFNLPLNDLQKLY"</p> <p>complement(join(2131..2980,3056..3446,3503..3816, 3875..3948,4063..4155,4209..4398,4452..4511,4569..4966, 5071..5175))</p> <p>/gene="CG3499"</p> <p>/product="CT1178"</p> <p>/db_xref="FLYBASE:FBan0003499"</p> <p>/db_xref="FLYBASE:FBgn0034792"</p> <p>complement(<2131..>5175)</p> <p>/gene="CG3499"</p> <p>/map="59B2-59B3"</p> <p>/db_xref="FLYBASE:FBan0003499"</p> <p>/db_xref="FLYBASE:FBgn0034792"</p> <p>complement(join(2317..2980,3056..3446,3503..3816, 3875..3948,4063..4155,4209..4398,4452..4511,4569..4966, 5071..5097))</p> <p>/gene="CG3499"</p> <p>/note="CG3499 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003499"</p> <p>/db_xref="FLYBASE:FBgn0034792"</p> <p>/protein_id="AAF46922.1"</p> <p>/db_xref="GI:7291497"</p> <p>/translation="MFSTHTSVSYLYLGNFSRPHYYSVNRTLKHSAGARLSKST STSRSHDLVDLNLRSASIQGMVERAARLNGILDRRLVDVLAQVTSMLPSMR DVRTLEESATQIGRVQLQNYQFEVSLTAGSVPTGANVKVIPTIPGLRLPLFSQO QLNQIRGFKTDRSEAEQKRNPTSRLLKALANSQORLDGTPLOAEKILRLLAKE EHSFNKSLKIAFAEYGLAAANSESPKSKTWKLYKTIIVIVFELGIFLSPFTTN GSVFRIOGNQVDEIINVTTFEDVKDEAKOELVEVEFLASPKFSNLGKLK GVLYGPPGKTLILARAVAGEAKVPFFHAGPFEDEVGQGARRVDFLFAAKARA PCVIFIDEISVGAKTNSVLHPYANTINQLSEMDFHONAGVIVLGATNRRDDL QALLRPFQVEVMVSPDEFTRKEILSLYTLKHDEIDLMLARGTSGTGGADLEN MINQAARAAIDGAETVSMKHLETARDKVLGMPERKARLPDEEANTITAYHEGGHATV AFYTKESHPLHKVTIMPRGPSLGHATVPEKERYHTVKAOLLAMDMTMMGSAEELV FGTDKITSGASSDLKQATSIATHVRDGMKSDKVLRTIEASKGLGTGDTLGPNTIEA VDABIKRILSDSYERAKAILKHTREHKAALAEALUYETLDLDDADIKAILNESQT"</p> <p>join(<5798..5857,5914..6403,6465..>6688)</p> <p>/gene="CG13533"</p> <p>/product="CT32908"</p> <p>/db_xref="FLYBASE:FBan0013533"</p> <p>/db_xref="FLYBASE:FBgn0034793"</p> <p>/evidence=not_experimental</p> <p><5798..>6688</p> <p>/gene="CG13533"</p> <p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0013533"</p> <p>/db_xref="FLYBASE:FBgn0034793"</p> <p>/evidence=not_experimental</p> <p>join(5798..5857,5914..6403,6465..6688)</p> <p>/gene="CG13533"</p> <p>/note="CG13533 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0013533"</p> <p>/db_xref="FLYBASE:FBgn0034793"</p> <p>/evidence=not_experimental</p> <p>/protein_id="AAF46923.1"</p> <p>/db_xref="GI:7291498"</p> <p>/translation="MDSPLNDGSHHPPHAPHPLADYQFSAEEVKALRECNTESSFOR SLPFGTGLGLLAYFGVNGVYLQGHVKYGNVPKVMGVILGYFVGKFSYQKCAEKMR</p>	CDS	<p>/map="59B2-59B3"</p> <p>/db_xref="FLYBASE:FBan0003501"</p> <p>/db_xref="FLYBASE:FBgn0034791"</p> <p>complement(join(662..1075,1557..1754))</p> <p>/gene="CG3501"</p> <p>/note="CG3501 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003501"</p> <p>/db_xref="FLYBASE:FBgn0034791"</p> <p>/protein_id="AAF46921.1"</p> <p>/db_xref="GI:7291496"</p> <p>/translation="MCDYKVSERAYAKLIFHAQYPHQAVNGLLLAEKTSKGSOVEIV DAIPLHQCLYVTPMAEVALMLIDAHAEERGLVIAGYAAAPENYDNOVQKTPAAKIA DKIQENKACFFVVDKMLTLQHDRAAIQVFNCPGSGARWSKAKETLQASDTLEG VSLLUKKGAMRDLDFONHLDNDPKNTDFNLPLNDLQKLY"</p> <p>complement(join(2131..2980,3056..3446,3503..3816, 3875..3948,4063..4155,4209..4398,4452..4511,4569..4966, 5071..5175))</p> <p>/gene="CG3499"</p> <p>/product="CT1178"</p> <p>/db_xref="FLYBASE:FBan0003499"</p> <p>/db_xref="FLYBASE:FBgn0034792"</p> <p>complement(<2131..>5175)</p> <p>/gene="CG3499"</p> <p>/map="59B2-59B3"</p> <p>/db_xref="FLYBASE:FBan0003499"</p> <p>/db_xref="FLYBASE:FBgn0034792"</p> <p>complement(join(2317..2980,3056..3446,3503..3816, 3875..3948,4063..4155,4209..4398,4452..4511,4569..4966, 5071..5097))</p> <p>/gene="CG3499"</p> <p>/note="CG3499 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003499"</p> <p>/db_xref="FLYBASE:FBgn0034792"</p> <p>/protein_id="AAF46922.1"</p> <p>/db_xref="GI:7291497"</p> <p>/translation="MFSTHTSVSYLYLGNFSRPHYYSVNRTLKHSAGARLSKST STSRSHDLVDLNLRSASIQGMVERAARLNGILDRRLVDVLAQVTSMLPSMR DVRTLEESATQIGRVQLQNYQFEVSLTAGSVPTGANVKVIPTIPGLRLPLFSQO QLNQIRGFKTDRSEAEQKRNPTSRLLKALANSQORLDGTPLOAEKILRLLAKE EHSFNKSLKIAFAEYGLAAANSESPKSKTWKLYKTIIVIVFELGIFLSPFTTN GSVFRIOGNQVDEIINVTTFEDVKDEAKOELVEVEFLASPKFSNLGKLK GVLYGPPGKTLILARAVAGEAKVPFFHAGPFEDEVGQGARRVDFLFAAKARA PCVIFIDEISVGAKTNSVLHPYANTINQLSEMDFHONAGVIVLGATNRRDDL QALLRPFQVEVMVSPDEFTRKEILSLYTLKHDEIDLMLARGTSGTGGADLEN MINQAARAAIDGAETVSMKHLETARDKVLGMPERKARLPDEEANTITAYHEGGHATV AFYTKESHPLHKVTIMPRGPSLGHATVPEKERYHTVKAOLLAMDMTMMGSAEELV FGTDKITSGASSDLKQATSIATHVRDGMKSDKVLRTIEASKGLGTGDTLGPNTIEA VDABIKRILSDSYERAKAILKHTREHKAALAEALUYETLDLDDADIKAILNESQT"</p> <p>join(<5798..5857,5914..6403,6465..>6688)</p> <p>/gene="CG13533"</p> <p>/product="CT32908"</p> <p>/db_xref="FLYBASE:FBan0013533"</p> <p>/db_xref="FLYBASE:FBgn0034793"</p> <p>/evidence=not_experimental</p> <p><5798..>6688</p> <p>/gene="CG13533"</p> <p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0013533"</p> <p>/db_xref="FLYBASE:FBgn0034793"</p> <p>/evidence=not_experimental</p> <p>join(5798..5857,5914..6403,6465..6688)</p> <p>/gene="CG13533"</p> <p>/note="CG13533 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0013533"</p> <p>/db_xref="FLYBASE:FBgn0034793"</p> <p>/evidence=not_experimental</p> <p>/protein_id="AAF46923.1"</p> <p>/db_xref="GI:7291498"</p> <p>/translation="MDSPLNDGSHHPPHAPHPLADYQFSAEEVKALRECNTESSFOR SLPFGTGLGLLAYFGVNGVYLQGHVKYGNVPKVMGVILGYFVGKFSYQKCAEKMR</p>	CDS	<p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0003495"</p> <p>/db_xref="FLYBASE:FBgn0034794"</p> <p>complement(<6781..>7947)</p> <p>/gene="CG3495"</p> <p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0003495"</p> <p>/db_xref="FLYBASE:FBgn0034794"</p> <p>complement(join(6781..6948,7005..7562,7621..7735, 7790..7914))</p> <p>/gene="CG3495"</p> <p>/note="CG3495 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003495"</p> <p>/db_xref="FLYBASE:FBgn0034794"</p> <p>/protein_id="AAF46924.1"</p> <p>/db_xref="GI:7291499"</p> <p>/translation="MKKVLVTGVTGLVKALEAVIKESPEDEQWFAGSKDADLTNL AATQALFAREKPTHVHILAMVGGSLFHNMMNLDLRLNLLINDVNLQTAHEQCCKV VSLCTCIFPDKTSYPIDETVMVHNGPPHPSNYGYSYAKRLIDVQNHAYHDKYGRVYTS VLCNIFGPHDNYNPEVSHVPIPMYIRMHQLVTKETDVPENDKVTTFVFGSMGRLQRFV YSRDLAELMWLNRVSEVPEIILSADEVQEVTFEVAQAVAKAFNNGRLVCDTSSKS DQYKKTASAKLSRSLFDYAFDTQLETAINASVWVYTENYDQAK"</p> <p>join(<8169..8207,8270..8363,8418..9382,9441..9527, 9583..9785,9838..10214,10275..10449,10506..10994, 11053..11379,11439..11662,11721..11803,11860..12007, 12077..12734,12796..12936,12995..>13304)</p> <p>/gene="CG3695"</p> <p>/product="CT12373"</p> <p>/db_xref="FLYBASE:FBan0003695"</p> <p>/db_xref="FLYBASE:FBgn0034795"</p> <p><8169..>13304</p> <p>/gene="CG3695"</p> <p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0003695"</p> <p>/db_xref="FLYBASE:FBgn0034795"</p> <p>join(8169..8207,8270..8363,8418..9382,9441..9527, 9583..9785,9838..10214,10275..10449,10506..10994, 11053..11379,11439..11662,11721..11803,11860..12007, 12077..12734,12796..12936,12995..13304)</p> <p>/gene="CG3695"</p> <p>/note="CG3695 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003695"</p> <p>/db_xref="FLYBASE:FBgn0034795"</p> <p>/protein_id="AAF46925.1"</p> <p>/db_xref="GI:7291500"</p> <p>/translation="METQVIDTVNEFLKVDLSLDAFVSIVYFKNTGEERATRFANDL VTAFGNVAENREQVLLYLRAAGASGHIKVLMAALVKLVDHIVTARMMLCDKVL CEKLDFAERTFWIESFLRIKRVIVQVDYKGYREIMKVCRDKAQFPLNVLNVTYMPQLL</p>	CDS	<p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0003495"</p> <p>/db_xref="FLYBASE:FBgn0034794"</p> <p>complement(<6781..>7947)</p> <p>/gene="CG3495"</p> <p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0003495"</p> <p>/db_xref="FLYBASE:FBgn0034794"</p> <p>complement(join(6781..6948,7005..7562,7621..7735, 7790..7914))</p> <p>/gene="CG3495"</p> <p>/note="CG3495 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003495"</p> <p>/db_xref="FLYBASE:FBgn0034794"</p> <p>/protein_id="AAF46924.1"</p> <p>/db_xref="GI:7291499"</p> <p>/translation="MKKVLVTGVTGLVKALEAVIKESPEDEQWFAGSKDADLTNL AATQALFAREKPTHVHILAMVGGSLFHNMMNLDLRLNLLINDVNLQTAHEQCCKV VSLCTCIFPDKTSYPIDETVMVHNGPPHPSNYGYSYAKRLIDVQNHAYHDKYGRVYTS VLCNIFGPHDNYNPEVSHVPIPMYIRMHQLVTKETDVPENDKVTTFVFGSMGRLQRFV YSRDLAELMWLNRVSEVPEIILSADEVQEVTFEVAQAVAKAFNNGRLVCDTSSKS DQYKKTASAKLSRSLFDYAFDTQLETAINASVWVYTENYDQAK"</p> <p>join(<8169..8207,8270..8363,8418..9382,9441..9527, 9583..9785,9838..10214,10275..10449,10506..10994, 11053..11379,11439..11662,11721..11803,11860..12007, 12077..12734,12796..12936,12995..>13304)</p> <p>/gene="CG3695"</p> <p>/product="CT12373"</p> <p>/db_xref="FLYBASE:FBan0003695"</p> <p>/db_xref="FLYBASE:FBgn0034795"</p> <p><8169..>13304</p> <p>/gene="CG3695"</p> <p>/map="59B2-59B4"</p> <p>/db_xref="FLYBASE:FBan0003695"</p> <p>/db_xref="FLYBASE:FBgn0034795"</p> <p>join(8169..8207,8270..8363,8418..9382,9441..9527, 9583..9785,9838..10214,10275..10449,10506..10994, 11053..11379,11439..11662,11721..11803,11860..12007, 12077..12734,12796..12936,12995..13304)</p> <p>/gene="CG3695"</p> <p>/note="CG3695 gene product"</p> <p>/codon_start=1</p> <p>/db_xref="FLYBASE:FBan0003695"</p> <p>/db_xref="FLYBASE:FBgn0034795"</p> <p>/protein_id="AAF46925.1"</p> <p>/db_xref="GI:7291500"</p> <p>/translation="METQVIDTVNEFLKVDLSLDAFVSIVYFKNTGEERATRFANDL VTAFGNVAENREQVLLYLRAAGASGHIKVLMAALVKLVDHIVTARMMLCDKVL CEKLDFAERTFWIESFLRIKRVIVQVDYKGYREIMKVCRDKAQFPLNVLNVTYMPQLL</p>	Query Match	58.1%; Score 18; DB 3; Length 299751;
Best Local Similarity	100.0%; Pred. No. 1.9;	Matches	18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	14	gtagcgctgagcaatttg	31						
Db	171879	GTAGCGCTGAGCAATTG	171862						
RESULT	9								
AF099920									
LOCUS	AF099920	34690 bp	DNA linear INV 16-NOV-2001						
DEFINITION	Caenorhabditis elegans cosmid H06H21, complete sequence.								
ACCESSION	AF099920								
VERSION	AF099920.2	GI:13559632							
KEYWORDS	HTG.								

gene	GRVVFCDGKQVCHIRGKLRKVKVINWGDIILVGLRDYODDKGDVILKYTPDEARR	
	LKNEGLIPENAKLNENDQDEGEVEFLDHVGDGAGCAKSDSDSDSKSDSEKSDSEAG	
CDS	SOKSEESDDSEEDSDSDNREEDLAAGRFKEDTRRGGNRGKNKYGKR*	
	5018..6996	
gene	/gene="H06H21.2"	
	Join(5018..5220,5624..5902,5950..6076,6154..6401,	
CDS	6775..6904,6955..6996)	
	/gene="H06H21.2"	
gene	/codon_start=1	
	/product="Hypothetical protein H06H21.2"	
CDS	/protein_id="AAK29844.1"	
	/db_xref="GI:13559635"	
gene	/translation="MSQAATCIHIVKIIYFSYFKLNEMIDVLLNLSAPSPPSY	
	DNATRIDLYFFLIAYEQVLAKIPPEPAPLSEFFYKLVINVLNDFRSTWLSVMA	
CDS	TIRLYLFSTRARRKKNFKPGFTLVIGLLSSIPLYIYFRIDVYKIGDWPMM	
	NCTFATSPESRFTVLQDGLGVKGYQLINGTSIKLPKPCFLPLVLTVMIL	
gene	VELKAKDHQRLNLTIRPVIINETSQVASTERTGLVVFITASTFLIEVPGNIVRL	
	QFGYDLAGYWRMATSVQAFSSAFVLHAALQCAIFELMSSQYRRRAVSKIFKNDRPPII	
CDS	IATVSSSY"	
	7392..10283	
gene	/gene="H06H21.1"	
	Join(7392..7546,7602..7820,8005..8283,8329..8455,	
CDS	9474..9727,10062..10191,10242..10283)	
	/gene="H06H21.1"	
gene	/codon_start=1	
	/product="Hypothetical protein H06H21.1"	
CDS	/protein_id="AAK29843.1"	
	/db_xref="GI:13559634"	
gene	/translation="MQVALIKNSYQVGNSSQKMDENTYNLATSAFFSYDNATKIQL	
	VHLKTVENFVPSIALIVFISLIIGVETTSFHLFVLSOKSLKSVILMMGVACDI	
CDS	LMAVASVFNYILYLTENFNAKCEPPLSLPPHYMYINIVINDFRGRTWLSVAMAL	
	IRWIVYKSTRAPKRVYLSFGAIFPGALALSPLFIYFRYDVIKIGDPLNN	
gene	CKSANFTVSNLYNLVQSEVYTANDGLGVKGYQLINGTSFKLIPCLPLDITLLIA	
	ELRKAKHQLMNSLTRLGNFNPETFSQGVASTERTGLVVFMTVTFVIEVPGGIVRV	
CDS	LQFGYDLAGYWRMATSVQAFSSAFVLHAALQCAIFELMSSQYRRRAVSKIFKNDRPPII	
	IATVSSSY"	
gene	complement(10489..14041)	
	/gene="H06H21.8"	
CDS	complement(join(10489..10683,10871..11130,11944..12067,	
	12543..12851,13608..13694,13850..14041))	
gene	/gene="H06H21.8"	
	/codon_start=1	
CDS	/product="Hypothetical protein H06H21.8"	
	/protein_id="AAK29842.1"	
gene	/db_xref="GI:13559633"	
	/translation="MKAKLQRLRVESFPFIEHFHKVDASFKELENKSAKSFWEIY	
CDS	VAHLKVGDGKVPSEVFKVPRISENVLRCEDESANHLNDVLLYYSKKENLYKHF	
	BYGSIPIPFKVPYETEDINGEATGTGIVAENLSEKVAFAVEHIPGLKHQILKMEALA	
gene	GLHSFLMKRDDKSYVESFVGAHGRETFSEGMQNMFEALTLENVSEVFCNDIRN	
	VIATVSSSY"	
Query Match	54.8%; Score 17; DB 3; Length 34690;	
	Best Local Similarity 100.0%; Pred. No. 9.1;	
Matches	17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	QY 1 cggtagctactcaggtag 17	
Db		
	26826 CGGTACTCTACGCTAG 26842	
RESULT 10	AC044838	
	AC044838	
LOCUS	178881 bp DNA linear HTG 09-MAY-2001	
	Mus musculus chromosome 11 clone RP23-65B3 map 11, WORKING DRAFT	
DEFINITION	SEQUENCE, 6 unordered pieces.	
	AC044838	
ACCESSION	AC044838.3	
	GI:12704668	
VERSION	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.	
	house mouse.	
KEYWORDS	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
SOURCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
	1 (bases 1 to 178881)	
REFERENCE	Birren,B., Linton,L., Nusbaum,C. and Lander,E.	
	1..178881	

TITLE
JOURNAL
REFERENCE
AUTHORS

Mus musculus chromosome 11, clone RP23-65B3
Unpublished
2 (bases 1 to 178881)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boquslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldin,J., Menesius,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT

Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 7, 2001 this sequence version replaced gi:9369528.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project information
Center project name: L8884
Center clone name: 65_B_3
----- Summary Statistics
Sequencing vector: M13; M77815; 40% of reads
Sequencing vector: plasmid; N/A; 60% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 176862 bases at least Q40
Consensus quality: 177787 bases at least Q30
Consensus quality: 178109 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 178381; sum-of-contigs
Quality coverage: 11.5 in Q20 bases; agarose-fp
Quality coverage: 10.9 in Q20 b.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 24550: contig of 24550 bp in length
* 24551 24650: gap of 100 bp
* 24651 39560: contig of 14910 bp in length
* 39561 39660: gap of 100 bp
* 39661 53388: contig of 13728 bp in length
* 53389 53488: gap of 100 bp
* 53489 85058: contig of 31570 bp in length
* 85059 85158: gap of 100 bp
* 85159 129894: contig of 44736 bp in length
* 129895 129994: gap of 100 bp
* 129995 178881: contig of 48887 bp in length.

Location/Qualifiers
1..178881
/organism="Mus musculus"

```

/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-65B3"
/clone_lib="rPCI-23 Female Mouse BAC"
misc_feature
1..24550
  note="assembly_fragment"
  clone_end:SP6
  vector_side:left
24551..39560
  note="assembly_fragment"
39561..53388
  note="assembly_fragment"
53489..85058
  note="assembly_fragment"
85159..129894
  note="assembly_fragment"
12995..178881
  note="assembly_fragment"
  clone_end:T7
  vector_side:right
48194 a 40622 c 40173 g 49390 t 502 others
ORIGIN

Query Match      54.8%; Score 17; DB 2; Length 178881;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtacctatcaggtagc 18
|||||
Db 40278 GGTACCTATCAGGTAGC 40294

RESULT 11
AL606531
LOCUS      AL606531      205165 bp      DNA      linear      HTG 30-JAN-2002
DEFINITION Mus musculus chromosome 11 clone RP23-104H11, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION  AL606531
VERSION     AL606531.6 GI:18135124
KEYWORDS   HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  1 (sites)
AUTHORS    Sims,S.
TITLE      Direct Submission
JOURNAL    Submitted (29-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 11, 2002 this sequence version replaced gi:17976609.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BM104H11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 204106 bases at least Q40
Consensus quality: 204442 bases at least Q30
Consensus quality: 204585 bases at least Q20
Insert size: 204665; sum-of-contigs
Insert size: 196989; 13.3% error; agarose-fp
Quality coverage: 8.71x in Q20 bases; sum-of-contigs Quality
coverage: 9.05x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
FEATURES
source
1..205165
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /chromosome="11"
  /clone="RP23-104H11"
  /clone_lib="rPCI-23"
misc_feature
1..16702
  note="assembly_fragment:01528"
  fragment_chain:1
  clone_end:SP6
  vector_side:left
16803..73634
  note="assembly_fragment:04543"
  fragment_chain:1
  fragment_chain:1
73735..115816
  note="assembly_fragment:00970"
  fragment_chain:1
115917..119281
  note="assembly_fragment:03278"
  fragment_chain:1
119382..182835
  note="assembly_fragment:04560"
  fragment_chain:1
182936..205165
  note="assembly_fragment:01498"
  fragment_chain:1
  clone_end:T7
  vector_side:right
55224 a 48879 c 47670 g 52892 t 500 others
ORIGIN

Query Match      54.8%; Score 17; DB 2; Length 205165;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtacctatcaggtagc 18
|||||
Db 145499 GGTACCTATCAGGTAGC 145515

RESULT 12
AC006719
LOCUS      AC006719      209365 bp      DNA      linear      HTG 23-FEB-1999
DEFINITION Caenorhabditis elegans clone Y1769, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
ACCESSION  AC006719
VERSION     AC006719.1 GI:4263219
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Caenorhabditis elegans.
ORGANISM   Caenorhabditis elegans
REFERENCE  1 (bases 1 to 209365)
AUTHORS    Waterston,R.H.
TITLE      The sequence of Caenorhabditis elegans clone
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 209365)
AUTHORS    Waterston,R.H.
TITLE      Direct Submission
JOURNAL    Submitted (23-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```


PUBMED 11053738
REFERENCE 2 (bases 1 to 1019)
AUTHORS Stubner S.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Stubner S., Biogeochemie,
Max-Planck-Institut f. terrestrische Mikrobiologie,
Karl-von-Frisch-Strasse, 35043 Marburg, GERMANY
FEATURES
source Location/Qualifiers
1. .1019
/organism="Desulfotomaculum sp. DEM-Kme99-1"
/db_xref="taxon:118333"
/clone="DEM-Kme99-1"
/note="isolated from rice paddy soil"
1. .1019
/gene="16S rRNA"
/product="16S ribosomal RNA"
1. .1019
/gene="16S rRNA"
/gene="16S rRNA"
BASE COUNT 257 a 228 c 339 g 194 t 1 others
ORIGIN
Query Match 51.6%; Score 16; DB 1; Length 1019;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 caggtagcgcgtgagca 26
|||||
Db 986 CAGGTAGCGCTGAGCA 1001
RESULT 15
HSU34962/c
LOCUS
DEFINITION Human transcription factor HCSX (hcsx) mRNA, complete cds.
ACCESSION U34962
VERSION U34962.1 GI:1314280
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1585)
AUTHORS Turbay, D., Wechsler, S.B., Blanchard, K.M. and Izumo, S.
TITLE Molecular cloning, chromosomal mapping, and characterization of the
human cardiac-specific homeobox gene hCsx
JOURNAL Mol. Med. 2 (1), 86-96 (1996)
MEDLINE 97056197
REFERENCE 2 (bases 1 to 1585)
AUTHORS Turbay, D.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-1995) David Turbay, Cardiovascular Research
Center, Department of Internal Medicine, The University of Michigan
Medical Center, 7220 MSRB III, 1150 West Medical Center Drive, Ann
Arbor, MI 48109, USA
FEATURES
source Location/Qualifiers
1. .1585
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q35"
/clone="pHcsx 1313"
/tissue_type="heart"
/clone_lib="lambda gt10 (Tamkun, et al., FASEB J. 5,
331-337)"
/dev_stage="adult"
1. .1585
/gene="hCsx"
177. .1151
/gene="hCsx"
/function="transcription factor"
/note="homeodomain containing protein"
/codon_start=1

/product="HCSX"
/protein_id="AAC50470.1"
/db_xref="GI:1314281"
/translation="MFPSPALTPFESVKDILNLEQQORSLAAAGELSARLEATLAPS
SCMLAAFKPEAVAGPEAAAGLPGLPELRAELGRAPSPAKCASAFPAAPAFYPRAYSDPDP
AKDPRAEKKELCALOKAVELEKTEADNAERPRARRRKRPRVLSQAQVYELERRFKQQ
RYLSAPERDQLASVLTSTQVKINWFQNRRYKCKRQRQDTLELGLPPPPPPARRI
AVPVLVRDQKPCIGDSAPYAYGVLNPGYGNAYPAYPGYGAACSPGYSCTAAYPA
GPSAPQAPATAAANNFVNFVGDLNAVQSPGIPQSNVSTLHGIRAW"
588. .768
/gene="hCsx"
/note="encodes helix-turn-helix motif"
misc_feature 256 a 569 c 497 g 263 t
BASE COUNT
ORIGIN
Query Match 51.6%; Score 16; DB 9; Length 1585;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 caggtagcgcgtgagca 26
|||||
Db 43 CAGGTAGCGCTGAGCA 28
Search completed: August 31, 2002, 22:35:31
Job time: 25160 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:51:27 ; Search time 824.51 Seconds
(without alignments)
64.553 Million cell updates/sec

Title: US-09-810-861b-2
Perfect score: 31
Sequence: 1 cggtaactatcaggtagcgtgagcaatttg 31

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	31	100.0	31	24	AAS17546	Human acetylcholin
C 2	31	100.0	5767	24	AAS17547	Plasmid pTM034. S
C 3	31	100.0	14446	24	AAS17548	Plasmid pTM036. S
C 4	16	51.6	591	21	AAF08167	Fusarium venenatum
C 5	16	51.6	6751	22	AAS09962	Genomic DNA #2 enc
6	16	51.6	16075	20	AAV99811	Gum gene cluster.
7	16	51.6	16079	8	AAW0753	Sequence of segmen
8	15	48.4	971	22	AAI69821	Shiitake mushroom
C 9	15	48.4	1725	24	AAS17549	Synthetic human ac

C 10	15	48.4	1845	24	AAS17493	Human cDNA encodin
11	15	48.4	2050	22	AAI69822	Shiitake mushroom
12	15	48.4	2051	19	AAV38497	Nucleic acid of th
C 13	15	48.4	2089	24	ABI199790	Mouse ischaemic co
C 14	15	48.4	2253	11	AAQ05998	Sequence encoding
C 15	15	48.4	2256	16	AAQ99002	Human acetylcholin
C 16	15	48.4	6290	22	AAI69823	Shiitake mushroom
C 17	15	48.4	12113	24	AAI17492	Human acetylcholin
C 18	14	45.2	358	20	AAI41139	Human secreted pro
C 19	14	45.2	393	20	AAI40564	Human secreted pro
C 20	14	45.2	588	23	AAI74531	DNA encoding novel
C 21	14	45.2	900	19	AAV63933	Mycobacterium tube
C 22	14	45.2	900	20	AAI81040	Nucleotide sequenc
C 23	14	45.2	1105	21	AAI52511	Arabidopsis thalia
C 24	14	45.2	1269	23	AAI74534	DNA encoding novel
C 25	14	45.2	1707	23	AAI94180	DNA encoding novel
C 26	14	45.2	1776	23	AAI74532	DNA encoding novel
C 27	14	45.2	1907	19	AAV38355	Transcription fact
C 28	14	45.2	1907	21	AAI36707	Nucleotide sequenc
C 29	14	45.2	1908	23	AAI94225	DNA encoding novel
C 30	14	45.2	1910	23	AAI74535	DNA encoding novel
C 31	14	45.2	1920	19	AAV40701	Human ELF3 protein
C 32	14	45.2	1996	21	AAI77813	Human cancer assoc
C 33	14	45.2	2112	23	ABL12431	Drosophila melanog
C 34	14	45.2	2353	23	AAI83049	DNA encoding novel
C 35	14	45.2	2379	23	ABL06673	Drosophila melanog
C 36	14	45.2	2374	23	AAI86684	DNA encoding novel
C 37	14	45.2	3774	23	AAI86684	DNA encoding novel
C 38	14	45.2	3774	23	AAI90257	DNA encoding novel
C 39	14	45.2	3776	23	AAI94279	DNA encoding novel
C 40	14	45.2	4233	23	ABL12430	Drosophila melanog
C 41	14	45.2	4543	23	ABL06672	Drosophila melanog
C 42	14	45.2	5679	23	ABL03552	Drosophila melanog
C 43	14	45.2	5823	23	AAI89361	DNA encoding novel
C 44	14	45.2	5828	23	ABL13644	Drosophila melanog
C 45	14	45.2	6048	19	AAV09029	Human hhl sodium c

ALIGNMENTS

RESULT 1
AAS17546
ID AAS17546 standard; DNA; 31 BP.

XX AAS17546;

XX 25-FEB-2002 (first entry)

XX Human acetylcholinesterase, PCR primer AChE-Kpn.

XX Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ss.

XX Homo sapiens.

OS Synthetic.

XX WO200171014-A2.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08468.

XX 17-MAR-2000; 2000US-190440P.

XX (MORT/) MOR T.

XX (SORE/) SOREQ H.

XX (ARNT/) ARNTZEN C.

XX (MASO/) MASON H.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H;
 XX WPI; 2002-055120/07.
 XX
 XX Production of a transgenic plant which contains a polynucleotide that
 PT encodes a human acetylcholinesterase which upon purification is
 PT effective against acetylcholinesterase poisoning -
 XX
 XX Claim 11; Page 4; 42pp; English.
 XX
 XX The invention relates to a method of producing a transgenic plant which
 CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
 CC which upon purification is effective against acetylcholinesterase
 CC poisoning. The method is used for treating a victim of
 CC acetylcholinesterase poisoning by administering a therapeutic amount of
 CC a physiologically active human acetylcholinesterase expressed in plant
 CC tissue. The extensive use of anticholinesterase pesticides with
 CC concurrent accidental poisoning, the threat of chemical warfare and
 CC environmental concerns demand the development of effective, inexpensive
 CC and stage countermeasures and bioremediation solutions. Prior art methods
 CC for treating AChE poisoning have used the muscarinic receptor antagonist
 CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
 CC The reversible carbamate, pyridostigmine bromide has also been used as a
 CC prophylactic. However, these conventional treatments have limited
 CC effectiveness and serious short and long-term side effects and may result
 CC in significant performance deficits and even permanent brain damage. This
 CC invention permits the utilisation of cholinesterases to counter-act the
 CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
 CC production of the enzymes is cost effective and the product is stable
 CC so that the injected enzymes have the advantage of having a long
 CC half-life. The transgenic form of the enzymes are also easy to purify.
 CC The present sequence represents PCR primer AChE-kpn used to amplify
 CC human AChE as described in the method of the invention.
 XX
 XX Sequence 31 BP; 7 A; 7 C; 9 G; 8 T; 0 other;
 SQ

Query Match 100.0%; Score 31; DB 24; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.4e-08;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggctacctatcaggtagcgctgagcaatttg 31
 |||||
 DB 1 cggctacctatcaggtagcgctgagcaatttg 31

RESULT 2
 AAS17547/C
 ID AAS17547 standard; DNA; 5767 BP.
 XX
 XX AAS17547;
 XX
 XX 25-FEB-2002 (first entry)
 XX
 XX Plasmid pTW034.
 XX
 XX Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
 KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
 KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
 KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTW034;
 KW circular; ds.
 XX
 XX Synthetic.
 OS
 XX WO200171014-A2.
 PN
 XX
 XX 27-SEP-2001.
 PD
 XX
 XX 16-MAR-2001; 2001WO-US08468.
 PF
 XX
 XX 17-MAR-2000; 2000US-190440P.
 PR
 XX
 XX (MORT/) MOR T.
 PA

PA (SORE/) SOREQ H.
 PA (ARNT/) ARNTZEN C.
 PA (MASO/) MASON H.
 PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
 XX
 XX Mor T, Soreq H, Arntzen C, Mason H;
 PI
 XX WPI; 2002-055120/07.
 DR
 XX
 XX Production of a transgenic plant which contains a polynucleotide that
 PT encodes a human acetylcholinesterase which upon purification is
 PT effective against acetylcholinesterase poisoning -
 XX
 XX Claim 11; Page 28-31; 42pp; English.
 XX
 XX The invention relates to a method of producing a transgenic plant which
 CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
 CC which upon purification is effective against acetylcholinesterase
 CC poisoning. The method is used for treating a victim of
 CC acetylcholinesterase poisoning by administering a therapeutic amount of
 CC a physiologically active human acetylcholinesterase expressed in plant
 CC tissue. The extensive use of anticholinesterase pesticides with
 CC concurrent accidental poisoning, the threat of chemical warfare and
 CC environmental concerns demand the development of effective, inexpensive
 CC and stage countermeasures and bioremediation solutions. Prior art methods
 CC for treating AChE poisoning have used the muscarinic receptor antagonist
 CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
 CC The reversible carbamate, pyridostigmine bromide has also been used as a
 CC prophylactic. However, these conventional treatments have limited
 CC effectiveness and serious short and long-term side effects and may result
 CC in significant performance deficits and even permanent brain damage. This
 CC invention permits the utilisation of cholinesterases to counter-act the
 CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
 CC production of the enzymes is cost effective and the product is stable
 CC so that the injected enzymes have the advantage of having a long
 CC half-life. The transgenic form of the enzymes are also easy to purify.
 CC The present sequence is plasmid pTW034, the pGPvkan derivative construct
 CC used in the generation of transgenic tomato plants that constitutively
 CC express human AChE.
 XX
 XX Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;
 SQ

Query Match 100.0%; Score 31; DB 24; Length 5767;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggctacctatcaggtagcgctgagcaatttg 31
 |||||
 DB 2571 CGGTACTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
 AAS17548/C
 ID AAS17548 standard; DNA; 14446 BP.
 XX
 XX AAS17548;
 XX
 XX 25-FEB-2002 (first entry)
 DT
 XX Plasmid pTW036.
 DE
 XX
 XX Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
 KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
 KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
 KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTW036;
 KW circular; ds.
 XX
 XX Synthetic.
 OS
 XX WO200171014-A2.
 PN
 XX
 XX 27-SEP-2001.
 PD

XX 16-MAR-2001; 2001WO-US08468.
 XX 17-MAR-2000; 2000US-190440P.
 XX (MORT/) MOR T.
 PA (SORE/) SOREQ H.
 PA (ARNT/) ARNTZEN C.
 PA (MASO/) MASON H.
 PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
 XX Mor T, SoREQ H, Arntzen C, Mason H;
 XX WPI; 2002-055120/07.
 XX
 XX Production of a transgenic plant which contains a polynucleotide that
 PT encodes a human acetylcholinesterase which upon purification is
 PT effective against acetylcholinesterase poisoning -
 XX
 XX Claim 11; Page 32-41; 42pp; English.
 XX
 CC The invention relates to a method of producing a transgenic plant which
 CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
 CC which upon purification is effective against acetylcholinesterase
 CC poisoning. The method is used for treating a victim of
 CC acetylcholinesterase poisoning by administering a therapeutic amount of
 CC a physiologically active human acetylcholinesterase expressed in plant
 CC tissue. The extensive use of anticholinesterase pesticides with
 CC concurrent accidental poisoning, the threat of chemical warfare and
 CC environmental concerns demand the development of effective, inexpensive
 CC and stage countermeasures and bioremediation solutions. Prior art methods
 CC for treating AChE poisoning have used the muscarinic receptor antagonist
 CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
 CC The reversible carbamate, pyridostigmine bromide has also been used as a
 CC prophylactic. However, these conventional treatments have limited
 CC effectiveness and serious short and long-term side effects and may result
 CC in significant performance deficits and even permanent brain damage. This
 CC invention permits the utilisation of cholinesterases to counter-act the
 CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
 CC production of the enzymes is cost effective and the product is stable
 CC so that the injected enzymes have the advantage of having a long
 CC half-life. The transgenic form of the enzymes are also easy to purify.
 CC The present sequence is plasmid pTM036, the pGPvkan derivative construct
 CC used in the generation of transgenic tomato plants that constitutively
 CC express human AChE.
 XX
 SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;

Query Match 100.0%; Score 31; DB 24; Length 14446;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtagctacaggtagcgctgagcaatttg 31
 |||||
 Db 1875 CGGTACCTATCAGTAGCGCTGAGCAATTG 1845

RESULT 4
 AAF08167/c
 ID AAF08167 standard; cDNA; 591 BP.
 XX AAF08167;
 XX
 DT 13-MAR-2001 (first entry)
 XX Fusarium venenatum EST SEQ ID NO:690.
 XX
 KW Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.

XX Fusarium venenatum.
 XX WO200056762-A2.
 XX 28-SEP-2000.
 XX 22-MAR-2000; 2000WO-US07781.
 XX 22-MAR-1999; 99US-0273623.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX (NOVO) NOVO NORDISK AS.
 XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 XX WPI; 2000-594572/56.
 XX Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags -
 XX
 XX Claim 86; Page 648; 3161pp; English.
 XX
 CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organisation of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
 CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
 CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
 CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
 CC all specifically claimed in the present invention.
 XX
 SQ Sequence 591 BP; 150 A; 154 C; 167 G; 119 T; 1 other;

Query Match 51.6%; Score 16; DB 21; Length 591;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcg 19
 |||||
 Db 474 TACCTATCAGGTAGCG 459

RESULT 5
 AAS09962/c
 ID AAS09962 standard; DNA; 6751 BP.
 XX AAS09962;
 XX
 DT 24-OCT-2001 (first entry)
 XX Genomic DNA #2 encoding human Csx/Nkx2.5.
 XX Csx/Nkx2.5; cardiac enhancer; cardiac cell; cardiomyocyte induction;
 KW therapeutic; heart tissue; gene therapy; human; ds.

```

XX OS Homo sapiens.
XX PN WO200151006-A2.
XX PD 19-JUL-2001.
XX PF 16-JAN-2001; 2001WO-US01511.
XX PR 14-JAN-2000; 2000US-0176419.
XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX PI Lee IW, Izumo S;
XX DR WPI; 2001-451809/48.
XX PT New cardiac specific cell enhancer elements, useful for specifically
XX PT expressing gene in cardiac cell, as earlier marker of cardiomyocyte
XX PT induction, e.g. for optimizing cardiomyocyte induction -
XX PS Claim 29; Fig 4B; 66pp; English.
XX CC The sequence represents the genomic sequence #2 of human Csx/Mx2.5. The
XX CC nucleic acid is useful for specifically expressing a gene in a cardiac
XX CC cell, as an earlier marker of cardiomyocyte induction, e.g. for
XX CC optimising cardiomyocyte induction. Genes expressed in the cardiac cell-
XX CC specific manner are useful for the targeted expression of genes encoding
XX CC therapeutic proteins for the treatment of damaged heart tissue. Cardiac
XX CC specific enhancer elements may be used for gene therapy.
XX SQ Sequence 6751 BP; 1459 A; 2008 C; 1753 G; 1490 T; 41 other;

Query Match 51.6%; Score 16; DB 22; Length 6751;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 caggtagcgtgagca 26
   |||||||
Db 2820 CAGGTAGCGCTGAGCA 2805

RESULT 6
ID AAV99811 standard; DNA; 16075 BP.
XX AC AAV99811;
XX DT 14-JUN-1999 (first entry)
XX DE Gum gene cluster.
XX KW Xanthan gum; gum; exopolysaccharide; gum gene cluster;
XX KW recombination; Xanthomonas; Sphingomonas; sphingans S88; sugar;
XX KW substrate; lactose; sucrose; starch; ss.
XX OS Xanthomonas campestris.
XX PN WO9856942-A1.
XX PD 17-DEC-1998.
XX PF 12-JUN-1998; 98WO-US12322.
XX PR 12-JUN-1998; 98US-0096942.
XX PR 12-JUN-1997; 97US-0049428.
XX PR 11-JUN-1998; 98US-0096867.
XX PA (SHIN-) SHINETSU BIO INC.
XX PA (SHIE ) SHINETSU CHEM CO LTD.
XX PI Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;

```

```

PI Yamazaki M;
XX DR WPI; 1999-080915/07.
XX PT Production of exopolysaccharide, e.g. xanthan gum by a Sphingomonas
XX PT species bacterium - containing the gumB-M genes of X. campestris,
XX PS and using lactose as a C source
XX PS Claim 8; Page 30-36; 50pp; English.
XX CC Xanthan gum is an exopolysaccharide produced by Xanthomonas species
XX CC and encoded by the gum gene cluster. The gum gene cluster of
XX CC Xanthomonas campestris can be inserted into a Sphingomonas species
XX CC of bacterium. The gum gene cluster is inserted into the Sphingomonas
XX CC S88 sps gene cluster which itself expresses an exopolysaccharide,
XX CC sphingans S88. The new novel recombinant Sphingomonas strain can then
XX CC produce xanthan gum from sugar substrates i.e. the cheese-making
XX CC by-product whey lactose, sucrose or starch. These are substrates
XX CC which the Xanthomonas campestris bacterium cannot utilise in its
XX CC production of Xanthan gum. Use of the Sphingomonas species also
XX CC minimises the presence of contaminating cellulases in the growth
XX CC medium into which the xanthan gum is secreted.
XX SQ Sequence 16075 BP; 2698 A; 4644 C; 5264 G; 3469 T; 0 other;

Query Match 51.6%; Score 16; DB 20; Length 16075;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 tcaggtagcgtgagc 25
   |||||||
Db 11351 tcaggtagcgtgagc 11366

RESULT 7
ID AAN70753 standard; DNA; 16079 BP.
XX AC AAN70753;
XX DT 20-MAY-1991 (first entry)
XX DE Sequence of segment of Xanthomonas campestris DNA that contains a
XX DE gene cluster that directs Xanthan biosynthesis.
XX KW Thickening agent; oil recovery; drilling fluid; ds.
XX OS Xanthomonas campestris.
XX PN WO8705938-A.
XX PD 08-OCT-1987.
XX PF 24-MAR-1987; 87WO-US00604.
XX PR 23-MAR-1987; 87US-0029530.
XX PR 24-MAR-1986; 86US-0842944.
XX PA (GETT-) GETTY SCI DEV CO.
XX CC Capage MA, Doherty DH, Betlach MR, Vanderslice RW;
XX PI WPI; 1987-291651/41.
XX DR Recombinant DNA prodn. of xanthan gum or its variants - by
XX PT transforming host cells with vector contg. DNA coding for enzymes
XX PT involved in polysaccharide synthesis
XX PS Example; Fig 10(1)-(39); 149pp; English.
XX CC Virtually all of the segment of Xanthomonas campestris DNA that
XX CC contains a gene cluster that directs Xanthan biosynthesis (AAN70753),

```

CC codes for protein products. Each gene is designated by a letter (see
 CC Fig 11) and its protein product is designated by that letter
 CC preceded by "gp" (AAP70455-67).
 XX
 SQ Sequence 16079 BP; 2707 A; 4649 C; 5261 G; 3462 T; 0 other;

Query Match 51.6%; Score 16; DB 8; Length 16079;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 tcaggtagcgcctgagc 25
 |||||

Db 11355 tcaggtagcgcctgagc 11370

RESULT 8

AAI69821
 ID AAI69821 standard; DNA; 971 BP.

XX AC AAI69821;

XX DT 18-DEC-2001 (first entry)

XX DE Shiitake mushroom tyrosinase gene terminator sequence.

XX KW Shiitake mushroom; tyrosinase; promoter; terminator;

XX KW transcription control; ds.

XX OS Lentinula edodes.

XX PN JP2001157586-A.

XX PD 12-JUN-2001.

XX PF 01-DEC-1999; 99JP-0342347.

XX PR 01-DEC-1999; 99JP-0342347.

XX PA (IWAT-) IWATE KEN.

XX WPI; 2001-599774/68.

XX New DNA which can function as a promoter for ordering the initiation of
 PT transcription in the tyrosinase gene of Lentinula edodes -

XX Claim 3; Page 13-14; 28pp; Japanese.

XX The invention relates to a DNA sequences that can be used to
 CC control the initiation of transcription of the tyrosinase gene of
 CC Lentinula edodes. The promoter region and terminator region of the
 CC Lentinula edodes tyrosinase gene were isolated. The DNA that can
 CC function as a promoter comprises a 2639 bp sequence given in the
 CC specification, or a sequence in which a base is deleted, replaced or
 CC added with respect to the 2639 bp sequence. The DNA that can function
 CC as a terminator comprises a 971 bp sequence given in the specification,
 CC or a sequence in which a base is deleted, replaced or added with respect
 CC to the 971 base pair sequence. The present sequence is a terminator
 CC region of the invention.

SQ Sequence 971 BP; 274 A; 204 C; 245 G; 248 T; 0 other;

Query Match 48.4%; Score 15; DB 22; Length 971;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 caggtagcgcctgagc 25

Db 83 caggtagcgcctgagc 97

RESULT 9

AAS17549/C
 ID AAS17549 standard; DNA; 1725 BP.

XX AC AAS17549;

XX DT 25-FEB-2002 (first entry)

XX DE Synthetic human acetylcholinesterase gene.

XX KW Human; acetylcholinesterase; AChE; antidote; pesticide;
 KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
 KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
 KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200171014-A2.

XX PD 27-SEP-2001.

XX PF 16-MAR-2001; 2001WO-US08468.

XX PR 17-MAR-2000; 2000US-190440P.

XX PA (MORT/) MOR T.

XX PA (SORE/) SOREQ H.

XX PA (ARNT/) ARNTZEN C.

XX PA (MASO/) MASON H.

XX PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX PI Mor T, SoREQ H, Arntzen C, Mason H;

XX DR WPI; 2002-055120/07.

XX Production of a transgenic plant which contains a polynucleotide that
 PT encodes a human acetylcholinesterase which upon purification is
 PT effective against acetylcholinesterase poisoning -

XX Claim 11; Page 41-42; 42pp; English.

XX The invention relates to a method of producing a transgenic plant which
 CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
 CC which upon purification is effective against acetylcholinesterase
 CC poisoning. The method is used for treating a victim of
 CC acetylcholinesterase poisoning by administering a therapeutic amount of
 CC a physiologically active human acetylcholinesterase expressed in plant
 CC tissue. The extensive use of anticholinesterase pesticides with
 CC concurrent accidental poisoning, the threat of chemical warfare and
 CC environmental concerns demand the development of effective, inexpensive
 CC and stage countermeasures and bioremediation solutions. Prior art methods
 CC for treating AChE poisoning have used the muscarinic receptor antagonist
 CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
 CC The reversible carbamate, pyridostigmine bromide has also been used as a
 CC prophylactic. However, these conventional treatments have limited
 CC effectiveness and serious short and long-term side effects and may result
 CC in significant performance deficits and even permanent brain damage. This
 CC invention permits the utilisation of cholinesterases to counter-act the
 CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
 CC production of the enzymes is cost effective and the product is stable
 CC so that the injected enzymes have the advantage of having a long
 CC half-life. The transgenic form of the enzymes are also easy to purify.
 CC The present sequence represents synthetic human AChE gene optimised for
 CC expression in plants.

SQ Sequence 1725 BP; 273 A; 553 C; 571 G; 328 T; 0 other;

Query Match 48.4%; Score 15; DB 24; Length 1725;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gcgctgagcaatttg 31

```

Db 1718 GCCTGAGCAATTG 1704
|||||
RESULT 10
AA17493/C
ID AAS17493 standard; cDNA; 1845 BP.
XX
XX AAS17493;
XX
XX 26-FEB-2002 (first entry)
XX
XX Human cDNA encoding acetylcholinesterase, ACHE.
XX
XX Human; ss; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;
KW genotyping; neurological disease; Parkinson's disease;
KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..1845
FT CDS
FT
FT /*tag= a
FT /product= "ACHE"
FT variation replace (36,T)
FT
FT /*tag= b
FT variation replace (101,A)
FT
FT /*tag= c
FT variation replace (183,C)
FT
FT /*tag= d
FT variation replace (838,G)
FT
FT /*tag= e
FT variation replace (873,T)
FT
FT /*tag= f
FT variation replace (1032,A)
FT
FT /*tag= g
FT variation replace (1057,A)
FT
FT /*tag= h
FT variation replace (1431,T)
FT
FT /*tag= i
FT variation replace (1057,A)
XX WO200179219-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US11853.
XX
XX 14-APR-2000; 2000US-197173P.
XX
XX (GENA-) GENAISANCE PHARM INC.
XX (KAZE/) KAZEMI A.
XX
XX Bentivegna SC, Chew A, Choi JY, Koshy B;
XX
XX WPI; 2002-055248/07.
XX P-PSDB; AAU11231.
XX
XX New polymorphic variants comprising acetylcholinesterase (ACHE)
XX isoenzyme, useful in expressing ACHE protein for use in screening for
XX candidate drugs to treat diseases related to ACHE activity, e.g.
XX neurological diseases or cancer -
XX
XX Claim 26; Fig 2; 79pp; English.
XX
XX The invention relates to a polynucleotide comprising a polymorphic

```

variant of an acetylcholinesterase (ACHE) gene or fragment, protein or complement, the variant comprising an ACHE isogene defined by a haplotype selected from haplotypes 1-20 listed in the specification. Also included are methods for haplotyping and genotyping the ACHE gene of an individual, a method for predicting a haplotype pair for the ACHE gene of an individual, a method for identifying an association between a trait and at least one haplotype or haplotype pair of ACHE gene, recombinant nonhuman organisms transformed or transfected with the polynucleotide where the organism expresses ACHE protein encoded by the first nucleotide sequence or encoded by the polymorphic variant sequence, an isolated antibody specific for and immunoreactive with ACHE, a method of screening for drugs targeting the polypeptide contacting ACHE polymorphic variant with a candidate agent and assaying for binding activity, a computer system for storing and analysing polymorphism data for ACHE gene and a genome anthology for ACHE gene which comprises ACHE isogenes defined by haplotypes 1-20 given in the specification. The polymorphisms are useful for studying the biological function of ACHE as well as in identifying drugs targeting this protein for the treatment of disorder related to its abnormal expression or function. The polymorphic variants may also be used in screening for compounds targeting ACHE to treat a specific condition or disease predicted to be associated with ACHE activity e.g. neurological diseases (e.g. Parkinson's disease and Alzheimer's disease), cancer, leukaemia, and tumours. The ACHE gene maps to human chromosome 7q22. The present sequence is the coding sequence of the ACHE gene.

Sequence 1845 BP; 302 A; 590 C; 608 G; 345 T; 0 other;

Query Match 48.4%; Score 15; DB 24; Length 1845;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcctgagcaatttg 31
|||||

Db 1718 GCCTGAGCAATTG 1704

RESULT 11
AA169822
ID AAI69822 standard; DNA; 2050 BP.
XX
XX AAI69822;
XX
DT 18-DEC-2001 (first entry)
XX
DE Shiitake mushroom tyrosinase gene.
XX
KW Shiitake mushroom; tyrosinase; promoter; terminator;
KW transcription control; ds.
XX
OS Lentinula edodes.
XX
PN JP2001157586-A.
XX
PD 12-JUN-2001.
XX
XX
PF 01-DEC-1999; 99JP-0342347.
XX
PR 01-DEC-1999; 99JP-0342347.
XX
PA (IWAT-) IWATE KEN.
XX
XX WPI; 2001-599774/68.
DR P-PSDB; AAG67089.
XX
XX
XX New DNA which can function as a promoter for ordering the initiation of
PT transcription in the tyrosinase gene of Lentinula edodes -
XX
XX Example 1; Page 14-17; 28pp; Japanese.
XX
XX The invention relates to a DNA sequences that can be used to
CC control the initiation of transcription of the tyrosinase gene of

CC Lentinula edodes. The promoter region and terminator region of the
 CC Lentinula edodes tyrosinase gene were isolated. The DNA that can
 CC function as a promoter comprises a 2639 bp sequence given in the
 CC specification, or a sequence in which a base is deleted, replaced or
 CC added with respect to the 2639 bp sequence. The DNA that can function
 CC as a terminator comprises a 971 bp sequence given in the specification,
 CC or a sequence in which a base is deleted, replaced or added with respect
 CC to the 971 base pair sequence. The present sequence is the
 CC tyrosinase gene of Lentinula edodes.
 XX
 SQ Sequence 2050 BP; 559 A; 509 C; 469 G; 513 T; 0 other;

Query Match 48.4%; Score 15; DB 22; Length 2050;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 caggtagcgctgagc 25
 |||||
 DB 1974 caggtagcgctgagc 1988

RESULT 12
 AAV38497
 ID AAV38497 standard; cDNA to mRNA; 2051 BP.

XX AAV38497;

DT 17-SEP-1998 (first entry)

DE Nucleic acid of the specification.

XX Shitake tyrosinase protein; Hokken number 57; ds.

KW Lentinus edodes.

FT Key Location/Qualifiers
 CDS 21..1904
 /*tag= a

PN JP10174586-A.

PD 30-JUN-1998.

XX 17-DEC-1996; 96JP-0337127.

PR 17-DEC-1996; 96JP-0337127.

PA (IWAT-) IWATE KEN.

XX WPI; 1998-421168/36.

DR N-PSDB; NAW62553.

XX Shitake tyrosinase gene - for producing large amounts of tyrosinase
 PT when introduced into various cells

XX Disclosure; Pages 7-10; 10pp; Japanese.

XX The present sequence appears in the specification, which describes
 CC a Shitake tyrosinase protein. The nucleic acid sequence was isolated
 CC from Shitake strain Hokken number 57. The gene can produce a large
 CC amount of tyrosinase when introduced to various cells.
 XX

SQ Sequence 2051 BP; 559 A; 512 C; 469 G; 511 T; 0 other;

Query Match 48.4%; Score 15; DB 19; Length 2051;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 caggtagcgctgagc 25
 |||||
 DB 1975 caggtagcgctgagc 1989

RESULT 13
 ABI99790/c

ID ABI99790 standard; cDNA; 2089 BP.

XX ABI99790;

DT 07-MAR-2002 (first entry)

XX Mouse ischaemic condition related cDNA sequence SEQ ID NO:890.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;

KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

XX Mus musculus.

XX WO200188188-A2.

PD 22-NOV-2001.

XX 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes -

XX Claim 2; Page 2244-2245; 2690pp; English.

XX The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC the expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.

SQ Sequence 2089 BP; 409 A; 638 C; 596 G; 446 T; 0 other;

Query Match 48.4%; Score 15; DB 24; Length 2089;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcgtgagcaatttg 31
 |||||

DB 1718 GCGCTGAGCAATTG 1704

RESULT 14
 AAQ05998/c

ID AAQ05998 standard; DNA; 2253 BP.

XX AAQ05998;

DT 16-JAN-1991 (first entry)

XX

DE Sequence encoding human acetylcholinesterase (hAChE).

XX Organophosphorous poisoning; OP; cancer; leukaemia;
 KW megakaryocytopoiesis; ovarian cancer; ds.
 XX Homo sapiens.
 OS
 PN EP388906-A.
 XX
 PD 26-SEP-1990.
 XX
 XX 20-MAR-1990; 90EP-0105274.
 PF
 XX 21-MAR-1989; 89IL-0089703.
 PR
 XX (YISS) YISSUM RES DEV CO.
 PA
 XX Soreq H, Zakut H;
 PI
 XX WPI; 1990-291865/39.
 DR
 XX P-PSDB; AAR06989.
 DR
 XX Human acetylcholinesterase DNA and prodn. of recombinant hAChE -
 PT for treatment of organo-phosphorous poisoning and diagnosis of haemo-
 PT cytopoietic disorders and ovarian carcinomas.
 PT
 XX Claim 5; Page 25; 47pp; English.
 PS
 XX Gene product is useful as an active pharmacological component for the
 CC prophylaxis and treatment of organophosphorous poisoning, and post-
 CC surgical apnea due to succinylcholine administration.
 CC CDNA probe to the sequence may be used in diagnosis of various
 CC leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
 XX
 XX Sequence 2253 BP; 390 A; 740 C; 694 G; 429 T; 0 other;

Query Match 48.4%; Score 15; DB 11; Length 2253;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctagagcaatttg 31
 |||||
 DB 1874 GCGCTGAGCAATTG 1860

RESULT 15
 AAQ99002/c
 ID AAQ99002 standard; DNA; 2256 BP.

XX AC AAQ99002;
 XX 31-MAR-1996 (first entry)
 XX
 DE Human acetylcholinesterase (AChE) gene.

XX Acetylcholinesterase; acetyl cholinesterase; EC-3.1.1.7;
 KW chromosome-7q22; acetylcholine-hydrolyzing enzyme; ss.
 XX

OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 160..2206
 FT /*tag= a

XX WO9523158-A1.

XX 31-AUG-1995.

XX 28-FEB-1995; 95WO-US02806.

XX 09-JAN-1995; 95US-0370156.

XX 28-FEB-1994; 94US-0202755.

XX

PA (KOHN/) KOHN K I.
 XX (YISS) YISSUM RES & DEV CO.

XX Shani M, Soreq H, Zakut H;

XX WPI; 1995-311499/40.

DR P-PSDB; AAR80726.

XX

PT Alternative forms of human acetyl cholinesterase (ChE) gene -
 PT expressed in transgenic animal assay system for evaluating anti-ChE
 PT activity of organo.phosphate(s), etc. or as model of ChE imbalance

XX Claim 3; Fig.1A; 55pp; English.

XX

CC This DNA sequence encoding human acetylcholinesterase is useful
 CC for producing transgenic animals which express AChE. The
 CC transgenic animals are in turn useful as an assay system for
 CC determining the anti-ChE activity of organophosphates, carbamates,
 CC anti-ChE drugs, plant glycoalkaloids and snake venoms. This gene
 CC contains a promoter region, 6 exons (E1 to E6) and 4 introns (I1-
 CC I4).

XX

SQ Sequence 2256 BP; 390 A; 757 C; 680 G; 429 T; 0 other;

Query Match 48.4%; Score 15; DB 16; Length 2256;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctagagcaatttg 31
 |||||

DB 1877 GCGCTGAGCAATTG 1863

Search completed: August 31, 2002, 22:51:31
 Job time: 16392 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 19:58:44 ; Search time 6064.54 Seconds
(without alignments)
68.992 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtaacctatcaggtagcgtgagcaatttg 31

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_etc.*
9: gb_est1.*
10: gb_est2.*
11: gb_etc.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	51.6	547	12	AZ586423
C 2	16	51.6	694	12	AZ274865
C 3	16	51.6	976	12	CNS02V61
C 4	16	51.6	1001	10	BI596784
C 5	15	48.4	52	9	AI205485
C 6	15	48.4	148	9	AW18501
C 7	15	48.4	177	10	BF892364
C 8	15	48.4	203	9	AW805755
C 9	15	48.4	214	10	BF875285
C 10	15	48.4	254	10	T28280
C 11	15	48.4	260	9	AW138046
C 12	15	48.4	292	9	AI205484
C 13	15	48.4	302	9	AA346027
C 14	15	48.4	348	9	BE146509
C 15	15	48.4	352	10	R28476
C 16	15	48.4	353	9	AI651502
C 17	15	48.4	353	9	AI652210

C 18	15	48.4	353	9	AW590331
C 19	15	48.4	359	10	N27052
C 20	15	48.4	367	9	AI341526
C 21	15	48.4	367	10	BE502360
C 22	15	48.4	378	10	F27586
C 23	15	48.4	380	9	AA933814
C 24	15	48.4	384	9	AI632562
C 25	15	48.4	384	9	AW612795
C 26	15	48.4	384	10	BE502241
C 27	15	48.4	384	10	BE504052
C 28	15	48.4	393	9	AA998511
C 29	15	48.4	395	9	AI766812
C 30	15	48.4	395	9	AI847791
C 31	15	48.4	407	10	BE551703
C 32	15	48.4	410	9	AI360141
C 33	15	48.4	411	9	AI690171
C 34	15	48.4	413	9	AA401251
C 35	15	48.4	420	9	AI825097
C 36	15	48.4	434	10	BI131084
C 37	15	48.4	442	9	BB733073
C 38	15	48.4	442	10	BG381540
C 39	15	48.4	443	10	BG381528
C 40	15	48.4	455	9	AW871154
C 41	15	48.4	467	9	AW828558
C 42	15	48.4	469	9	AI831696
C 43	15	48.4	476	12	AQ661569
C 44	15	48.4	479	9	AW783232
C 45	15	48.4	480	10	N36355

ALIGNMENTS

RESULT 1
AZ586423/c 547 bp DNA linear GSS 13-DEC-2000
LOCUS 1M0392M15F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
DEFINITION Clone UUGCIM0392M15 F, DNA sequence.

ACCESSION AZ586423
VERSION 1 GI:11708613

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 547)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Lonacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0392 row: M column: 15
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 547.
Location/Qualifiers
1. .547
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="Mouse 10kb plasmid UUGCIM library"

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydronynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gii4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      179 a 130 c 122 g 116 t
ORIGIN

```

```

Query Match      51.6%; Score 16; DB 12; Length 547;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 12 aggtacgcctgagcaa 27
      |||||
Db 304 AGGTAGCGCTGAGCAA 289

```

```

RESULT 2
LOCUS      A2274865/c
DEFINITION      A2274865      694 bp      DNA      linear      GSS 26-JUL-2000
      RPCI-23-118H10-TV RPCI-23 Mus musculus genomic clone RPCI-23-118H10

```

```

ACCESSION      A2274865
VERSION        A2274865.1 GI:9488482
KEYWORDS
SOURCE

```

```

ORGANISM

```

```

REFERENCE
AUTHORS      Zhao S., Nierman W., Feldblyum, T., Malek J., Shatsman, S., Akinet
      B., Levins, M., Mcgann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
      and Fraser, C.M.
      Mus musculus house mouse.

```

```

TITLE
JOURNAL

```

```

COMMENT      Mouse BAC End Sequences from Library RPCI-23
      Other GSSs: RPCI-23-118H10.TJ
      Department of Eukaryotic Genomics
      The Institute for Genomic Research
      9712 Medical Center Dr., Rockville, MD 20850, USA
      Tel: 301 838 0200
      Fax: 301 838 0208
      Email: szhac@tigr.org

```

```

      Clones are derived from the mouse BAC library RPCI-23. For BAC
      library availability, please contact Pieter de Jong
      (pieter@dejong.med.buffalo.edu). Clones may be purchased from
      BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
      or from Resea ch Genetics (info@resgen.com). BAC end page:
      http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html
      Plate: 118 row: H column: 10
      Seq primer: T7
      Class: BAC ends.

```

```

FEATURES
source
      Location/Qualifiers
      1..694
      /organism="Mus musculus"
      /strain="C57BL/6J"

```

```

/db_xref="taxon:10090"
/clone="RPCI-23-118H10"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/notes="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:
      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
      brain genomic DNA was isolated and partially digested
      with a combination of EcoRI and EcoRI Methylase. Size
      selected DNA was cloned into the pBACE3.6 vector at the
      EcoRI sites. The ligation products were transformed into
      DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      237 a 90 c 94 g 273 t
ORIGIN

```

```

Query Match      51.6%; Score 16; DB 12; Length 694;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3 gtacctatcagctagc 18
      |||||
Db 146 GTACCTATCAGGTAGC 131

```

```

RESULT 3
LOCUS      CNS02V61
DEFINITION      CNS02V61      976 bp      DNA      linear      GSS 15-MAY-2000
      Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
      173P07 of library G from Tetraodon nigroviridis, genomic survey
      sequence.

```

```

ACCESSION      AL215362
VERSION        AL215362.1 GI:7874181
KEYWORDS      GSS; genome survey sequence.
SOURCE

```

```

ORGANISM

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

```

```

REFERENCE
AUTHORS

```

```

Roest-Crollius H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.,
Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.

```

```

TITLE
JOURNAL

```

```

REFERENCE
AUTHORS

```

```

Roest-Crollius H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.

```

```

Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

```

```

TITLE
JOURNAL

```

```

REFERENCE
AUTHORS

```

```

TITLE
JOURNAL

```

```

COMMENT

```

```

This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.

```

```

FEATURES
source
      Location/Qualifiers
      1..976
      /organism="Tetraodon nigroviridis"
      /db_xref="taxon:99883"
      /clone_lib="G"
      /clone="173P07"
      /note="Genoscope sequence ID : COAG173CH04SP1-end ;
      PUC-Ori"

```

```

BASE COUNT      267 a 194 c 216 g 272 t 27 others
ORIGIN

```


JOURNAL
COMMENT

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
source

1. .148
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2874155"
/clone_lib="NCI-CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones IDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 17 a 72 c 31 g 28 t
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgcctgagcaatttcg 31
|||||
Db 15 GCGCTGAGCAATTTC 1

RESULT 7

BF892364
LOCUS BF892364 177 bp mRNA linear EST 18-JAN-2001
DEFINITION CM1-MT0102-041100-434-d07 MT0102 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF892364
VERSION BF892364.1 GI:12283823
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

FEATURES
source

1. .177
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0102"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 39 a 54 c 41 g 43 t
ORIGIN

Query Match 48.4%; Score 15; DB 10; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cggtaacctatcaggt 15
|||||
Db 61 CGTACCTATCAGGT 75

RESULT 8

AW805755
LOCUS AW805755 203 bp mRNA linear EST 16-MAY-2000
DEFINITION QV1-UM0105-180400-162-b06 UM0105 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW805755
VERSION AW805755.1 GI:7857625
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-QV1-UM0105-180 400-162-b06st3-2000-04-18st4-1)

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM1st2-CM1-MT0102-041100-434-d07st3-2000-11-04st4-1)

Seq primer: puc 18 forward

High quality sequence start: 5

High quality sequence stop: 177.

Seq primer: puc 18 forward
High quality sequence start: 35
High quality sequence stop: 203.

FEATURES

Location/Qualifiers
1. .203
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UM0105"
/dev_stage="Adult"
/note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 47 a 34 c 47 g 75 t
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtacctatcaggta 16
|||||
DB 43 GGTACTATCAGGTA 57

RESULT 9

LOCUS BF875285 214 bp mRNA linear EST 17-JAN-2001
DEFINITION QV3-ET0099-101100-390-e10 ET0099 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF875285
VERSION BF875285.1 GI:12265415
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 214)

DIAS Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bait, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE
PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV3&t2=QV3-ET0099-101100-390-e10&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 213.

FEATURES

source

Location/Qualifiers
1. .214
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0099"
/dev_stage="Adult"
/note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 48 a 34 c 45 g 87 t
ORIGIN

Query Match 48.4%; Score 15; DB 10; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 tacctatcaggtagc 18
|||||
DB 16 TACCTATCAGGTAGC 30

RESULT 10

T28280/c

LOCUS

DEFINITION T28280 254 bp mRNA linear EST 06-SEP-1995
acetylcholinesterase (HT:518), mRNA sequence.

ACCESSION T28280

VERSION T28280.1

KEYWORDS GI:610378

SOURCE EST.

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 254)

AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult

, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

, O., Sutton, G., Blake, J.A., Brandon, R.C., Chiu, M.-W., Clayton, R.A.,

Cline, R.T., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M.,

FitzHugh, W.M., Fritchman, J.L., Geoghagen, N.S.M., Glodok, A.,

Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, J.P., Kelley, J.M.,

Klimmek, K.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Mellegre, S.M.,

Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pellegrino, S.M.,

Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,

Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,

Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.-J.,

Dimke, D., Feng, P., Ferrie, A., Fischer, C., Hastings, G.A., He, W.-W.,

Hu, J.-S., Greene, J.M., Gruber, J., Hudson, P., Kim, A., Kozak, D.L.,

Kunsch, C., Ji, H., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei

, Y.-F., Wang, J., Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon

, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and

Venter, J.C.

Initial Assessment of Human Gene Diversity and Expression Patterns

Based Upon 83 Million Basepairs of cDNA Sequence

Nature 377, 3-174 (1995)

96026280

Other ESTs: THC20776

Contact: Venter, J.C.

The Institute for Genomic Research

932 Clopper Rd, Gaithersburg, MD 20878

Tel: 3018699056

Fax: 3018699423

Email: tdbinfo@tdb.tigr.org

For clone availability, additional sequence and expression

information related to this EST, please contact the TIGR Database

(tdbinfo@tdb.tigr.org)

Seq primer: M13 Reverse.

Location/Qualifiers

1. .254

/organism="Homo sapiens"

/db_xref="ATCC (inhost):101797"

/db_xref="taxon:9606"

/clone_lib="Human Embryo"

/tissue_type="embryo"

45 a 89 c 78 g 41 t 1 others

BASE COUNT

ORIGIN

Query Match 48.4%; Score 15; DB 10; Length 254;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcgtgagcaatttg 31
 |||||
 DB 101 GCGCTGAGCAATTG 87

RESULT 11
 AW138046/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AW138046 260 bp mRNA linear EST 29-OCT-1999
 UI-H-BII-abw-b-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
 IMAGE:2713294 3', mRNA sequence.

AW138046
 AW138046.1 GI:6142364
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 260)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-f@mail.nih.gov

Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: NCI-CGAP clone distribution
 Information can be found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/bbrp/image/image.html The following repetitive
 elements were found in this cDNA sequence: 147-199, >(GGA
)n#Simple_repeat
 Seq primer: M13 Forward
 POLYA=NO.

FEATURES
 source
 1. .260
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2713294"
 /clone_lib="NCI_CGAP_Sub3"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NCI_CGAP_Sub3 library is a subtracted library derived from
 the NCI_CGAP_Sub1 library, which is a subtracted library
 derived from BI. BI constitutes a mixture of 21
 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4
 , NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
 NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
 NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
 NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Le12,
 NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
 NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
 NCI_CGAP_Brn25. These 21 libraries were pooled and a
 single-stranded DNA preparation of the resulting mixture
 was used as a tracer in a subtractive hybridization with
 a driver whose composition is detailed below:
 NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
 3798-3803 (IMAGE Clones 1322376-1323911, 1456008-1456775
 , 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342
 , 3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831,
 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1
 LLAM 3573-3582, 3851-3854 (IMAGE Clones 1414920-1417991,
 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
 3716-3720, 3733-3735 (IMAGE Clones 1257096-1258631,
 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones
 985608-986759, 1101192-1101959, 1217928-1220615);

NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
 Clones 1057416-1061255, 1144584-1145351). Subtraction
 was performed as previously described (Bonaldo, Lennon &
 Soares (1996)): Normalization and Subtraction: Two
 Approaches To Facilitate Gene Discovery. Genome Research
 6, 791-806.

BASE COUNT 30 a 112 c 72 g 46 t
 ORIGIN TAG_SEQ=None found"

Query Match 48.4%; Score 15; DB 9; Length 260;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcgtgagcaatttg 31
 |||||
 DB 80 GCGCTGAGCAATTG 66

RESULT 12
 AI205484/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AI205484 292 bp mRNA linear EST 03-FEB-1999
 qr26a12.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:1941982 3'
 similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN
); contains element MSRI MSRI repetitive element ;, mRNA sequence.

AI205484
 AI205484.1 GI:3764156
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 292)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-f@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
[www-bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)
 Insert Length: 773 Std Error: 0.00
 Seq primer: -400P from Gibco
 High quality sequence stop: 235.

FEATURES
 source
 1. .292
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1941982"
 /clone_lib="NCI_CGAP_GC6"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
 from the normalized library NCI_CGAP_GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clones
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 45 a 123 c 72 g 52 t
 ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcagagcaattg 31
 |||||
 Db 50 GCGCTGAGCAATTG 36

RESULT 13
 AA346027/c
 LOCUS
 DEFINITION EST52141 Greater Omentum I Homo sapiens cDNA 5' end similar to
 acetylcholinesterase, mRNA sequence.
 ACCESSION AA346027
 VERSION AA346027.1 GI:1998286
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 302)

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S., Glodek,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
 , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
 ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT Other_ESTs: THC103848
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlavetigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse

FEATURES
 Location/Qualifiers
 1..302

/organism="Homo sapiens"
 /db_xref="ATCC (Inhost):147585"
 /db_xref="taxon:9606"
 /clone_lib="Greater omentum I"
 /tissue_type="omentum"
 /dev_stage="adult"
 /note="Organ: NULL; Vector: pBluescript SK-; Site_1: EcoRI
 ; Site_2: XhoI"

BASE COUNT 60 a 103 c 91 g 44 t 4 others
 ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcagagcaattg 31
 |||||
 Db 194 GCGCTGAGCAATTG 180

RESULT 14
 BE146509/c

LOCUS
 DEFINITION RCI-HT0217-151099-011-g06 HT0217 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE146509
 VERSION BE146509.1 GI:8609233
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bais,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=62-rc1-HT0217-151099-011-g06&t3=1999-10-15&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 8
 High quality sequence stop: 348.

FEATURES
 source

1..348
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0217"
 /dev_stage="Adult"
 /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 122 a 65 c 74 g 87 t
 ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 348;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtaccatcaggta 16
 |||||
 Db 82 GGTACCTATCAGGTA 68

RESULT 15
 R28476/c

LOCUS
 DEFINITION Yh50g12_r1 Soares placenta Nb2HP Homo sapiens cDNA clone
 IMAGE:133222 5', mRNA sequence.

ACCESSION R28476
 VERSION R28476.1 GI:784611
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 352)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilton RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1068
 High quality sequence stops: 273
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1068 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 273.
 FEATURES
 Location/Qualifiers
 1..352
 /organism="Homo sapiens"
 /db_xref="GBB:538912"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:133222"
 /clone_lib="Soares placenta Nb2Hp"
 /sex="Female"
 /dev_stage="placenta obtained at birth (full term)"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGCAACAATCCGCGCGCAGGAATTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 119 a 67 c 88 g 78 t
 ORIGIN
 Query Match 48.4%; Score 15; DB 10; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ggtaccctatcagcta 16
 |
 Db 108 GGTACCTATCAGGTA 94

Search completed: August 31, 2002, 19:58:51
 Job time: 16340 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:37:20 ; Search time 154.73 Seconds
(without alignments)
49.212 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtaactatcaggtagcgtgagcaatttg 31

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	51.6	16075	3	US-09-096-942-1
2	16	51.6	16075	3	US-09-096-867-1
C 3	15	48.4	1845	1	US-07-732-962A-1
C 4	15	48.4	1845	5	PCT-US92-06106-1
C 5	15	48.4	2256	2	US-08-318-826A-5
C 6	15	48.4	2256	2	US-08-370-156-1
C 7	15	48.4	2256	3	US-08-814-095-1
C 8	15	48.4	3016	2	US-08-318-826A-7
C 9	15	48.4	3016	2	US-08-370-156-5
C 10	15	48.4	3016	3	US-08-814-095-5
C 11	15	48.4	3096	2	US-08-318-826A-6
C 12	15	48.4	3096	2	US-08-370-156-3
C 13	15	48.4	3096	3	US-08-814-095-3
C 14	15	48.4	35060	3	US-08-814-095-7
C 15	14	45.2	867	4	US-09-273-839A-7
C 16	14	45.2	1920	1	US-08-746-789A-1
C 17	14	45.2	6048	4	US-09-634-920-3
C 18	14	45.2	17138	4	US-09-813-819-3
C 19	14	45.2	17138	4	US-09-920-048-3
C 20	13	41.9	289	2	US-08-967-101-23
C 21	13	41.9	289	2	US-08-592-541-23
C 22	13	41.9	289	3	US-09-124-698-23
C 23	13	41.9	289	4	US-09-127-480-23
C 24	13	41.9	289	4	US-08-496-841C-23
C 25	13	41.9	335	2	US-08-623-906A-12
C 26	13	41.9	401	3	US-08-581-918A-27
27	13	41.9	401	4	US-08-346-147B-27

28	13	41.9	404	3	US-08-581-918A-26	Sequence 26, Appl
29	13	41.9	404	4	US-08-346-147B-26	Sequence 26, Appl
30	13	41.9	837	4	US-09-570-367C-20	Sequence 20, Appl
31	13	41.9	837	4	US-09-570-367C-22	Sequence 22, Appl
32	13	41.9	1149	1	US-08-474-177-3	Sequence 3, Appl
33	13	41.9	1149	1	US-08-487-033-3	Sequence 3, Appl
34	13	41.9	1149	1	US-08-480-810-3	Sequence 3, Appl
35	13	41.9	1149	2	US-08-508-735-3	Sequence 3, Appl
36	13	41.9	1149	2	US-08-848-251-3	Sequence 3, Appl
37	13	41.9	1149	2	US-08-486-047-3	Sequence 3, Appl
38	13	41.9	1149	3	US-09-120-130-3	Sequence 3, Appl
39	13	41.9	1149	3	US-09-115-252-3	Sequence 3, Appl
40	13	41.9	1149	3	US-08-986-515-3	Sequence 3, Appl
41	13	41.9	1149	3	US-09-120-128-3	Sequence 3, Appl
42	13	41.9	1149	4	US-09-120-129-3	Sequence 3, Appl
43	13	41.9	1149	4	US-09-201-139-3	Sequence 3, Appl
44	13	41.9	1149	4	US-09-120-131-3	Sequence 3, Appl
C 45	13	41.9	1608	1	US-08-393-985-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-09-096-942-1
; Sequence 1, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,942
; EARLIER FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-942-1

Query Match 51.6%; Score 16; DB 3; Length 16075;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 tcaggtagcgtgagc 25

DB 11351 tcaggtagcgtgagc 11366

RESULT 2

US-09-096-867-1
; Sequence 1, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867

; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-867-1

Query Match 51.6%; Score 16; DB 3; Length 16075;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 tcaggtagcgctgagc 25
|||||
Db 11351 tcaggtagcgctgagc 11366

RESULT 3
US-07-732-962A-1/C
; Sequence 1, Application US/07732962A
; Patent No. 5248604
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
; TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 664-0525
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842

US-07-732-962A-1
Query Match 48.4%; Score 15; DB 1; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 gcgctgagcaatttg 31

Db 1718 GCGCTGAGCAATTG 1704
|||||

RESULT 4
PCT-US92-06106-1/c
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842

PCT-US92-06106-1
Query Match 48.4%; Score 15; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 gcgctgagcaatttg 31
|||||

Db 1718 GCGCTGAGCAATTG 1704
|||||

RESULT 5
US-08-318-826A-5/C
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410

US-08-318-826A-5/C
Query Match 48.4%; Score 15; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 gcgctgagcaatttg 31
|||||

Db 1718 GCGCTGAGCAATTG 1704
|||||

RESULT 5
US-08-318-826A-5/C
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410

US-08-318-826A-5/C
Query Match 48.4%; Score 15; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 gcgctgagcaatttg 31
|||||

Db 1718 GCGCTGAGCAATTG 1704
|||||

RESULT 5
US-08-318-826A-5/C
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410

US-08-318-826A-5/C
Query Match 48.4%; Score 15; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 gcgctgagcaatttg 31
|||||

Db 1718 GCGCTGAGCAATTG 1704
|||||

RESULT 5
US-08-318-826A-5/C
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410

US-08-318-826A-5/C
Query Match 48.4%; Score 15; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 gcgctgagcaatttg 31
|||||

Db 1718 GCGCTGAGCAATTG 1704
|||||

;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: US
;; ZIP: 48334
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/318,826A
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: 2391.00001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 539-5050
;; TELEFAX: (248) 539-5055
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: /note= "splice variant: Exons 1, 2,
;; OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match 48.4%; Score 15; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 6
US-08-370-156-1/c
;; Sequence 1, Application US/08370156
;; Patent No. 5932780
;; GENERAL INFORMATION:
;; APPLICANT: Soreq, Hermona
;; APPLICANT: Zakut, Haim
;; APPLICANT: Shani, Moshe
;; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
;; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
;; NUMBER OF SEQUENCES: 27
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Reising, Ethington, Barnard & Perry
;; STREET: P.O. Box 4390
;; CITY: Troy
;; STATE: Michigan
;; COUNTRY: US
;; ZIP: 48099
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/370,156
;; FILING DATE:
;; CLASSIFICATION: 536

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810) 689-3500
;; TELEFAX: (810) 689-4071
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-370-156-1

Query Match 48.4%; Score 15; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 7
US-08-814-095-1/c
;; Sequence 1, Application US/08814095
;; Patent No. 6025183
;; GENERAL INFORMATION:
;; APPLICANT: Soreq, Hermona
;; APPLICANT: Zakut, Haim
;; APPLICANT: Shani, Moshe
;; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
;; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: KOHN & ASSOCIATES
;; STREET: 30500 No. 6025183thwestern Highway, Suite 410
;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: U.S.
;; ZIP: 48334
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/814,095
;; FILING DATE:
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Montgomery, Ilene N.
;; REGISTRATION NUMBER: 38,972
;; REFERENCE/DOCKET NUMBER: 2391.00066
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 539-5050
;; TELEFAX: (248) 539-5055
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "ACHE gene comprising exons
;; DESCRIPTION: 2, 3, 4 and 6"
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
US-08-814-095-1

Query Match 48.4%; Score 15; DB 3; Length 2256;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||

Db 1877 GCGCTGAGCAATTG 1863

RESULT 8
US-08-318-826A-7/c
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; OTHER INFORMATION: /note= "Splice Variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and 6"
US-08-318-826A-7

Query Match 48.4%; Score 15; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||

Db 1877 GCGCTGAGCAATTG 1863

RESULT 9
US-08-370-156-5/c
; Sequence 5, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
US-08-370-156-5

Query Match 48.4%; Score 15; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||

Db 1877 GCGCTGAGCAATTG 1863

RESULT 10
US-08-814-095-5/c
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.

; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced Ache
; DESCRIPTION: comprising exons 2, 3, 4, 5 and 6"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; US-08-814-095-5

Query Match 48.4%; Score 15; DB 3; Length 3016;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 11
US-08-318-826A-6/c
; Sequence 6, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318.826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.

; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and the translated portion of Intron 4 (readthrough
US-08-318-826A-6

Query Match 48.4%; Score 15; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 12
US-08-370-156-3/c
; Sequence 3, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370.156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```
; ;
; NAME/KEY: CDS
; LOCATION: 160...1959
US-08-370-156-3

Query Match      48.4%; Score 15; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcagcaatttg 31
|||||

Db 1877 GCCTGAGCAATTG 1863

RESULT 13
US-08-814-095-3/c
; Sequence 3, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced AchE
; DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
; DESCRIPTION: of Intron 4 (readthrough)."
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160...1959
US-08-814-095-3

Query Match      48.4%; Score 15; DB 3; Length 3096;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgcagcaatttg 31
|||||

Db 1877 GCCTGAGCAATTG 1863

RESULT 14
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE
; DESCRIPTION: promoter, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 4089...22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465...22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090...25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "{translation start:
; OTHER INFORMATION: 24110}"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
```

```

, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 8
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (31894..32080)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 9
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (31363..31534)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 10
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (31131..31284)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 11
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (30816..31011)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 12
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (30470..30626)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 13
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (30187..30274)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 14
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (29945..30073)
, OTHER INFORMATION: /gene= "AR"
, OTHER INFORMATION: /number= 15
, FEATURE:
, NAME/KEY: exon
, LOCATION: complement (29664..29856)
, OTHER INFORMATION: /gene= "ARS"
, OTHER INFORMATION: /number= 16
US-08-814-095-7

Query Match 48.4%; Score 15;
Best Local Similarity 100.0%; Pred. No.
Matches 15; Conservative 0; Mismatchch

QY 17 gcgctgagcaatttg 31
|||||
Db 27168 GCGCTGAGCAATTG 27154

RESULT 15
US-09-273-839A-7
, Sequence 7, Application US/09273839A
, Patent No. 6329156
, GENERAL INFORMATION:
, APPLICANT: Cirino, Nick M
, APPLICANT: Jackson, Paul J
, APPLICANT: Lehnert, Bruce E
, TITLE OF INVENTION: Disruption of Anthrax
, FILE REFERENCE: S-89,662
, CURRENT APPLICATION NUMBER: US/09/273,839A
, CURRENT FILING DATE: 1999-03-22
, NUMBER OF SEQ ID NOS: 12
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 7
, LENGTH: 867
, TYPE: DNA
, ORGANISM: Bacillus anthracis

```

Qy 17 gcgctgagcaatttg 31
 |||||
Db 27168 GCGCTGAGCAATTG 27154

```

RESULT 15
US-09-273-839A-7
; Sequence 7, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Bacillus anthracis

```

Search completed: August 31, 2002, 22:37:24
Job time: 19308 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:35:31 ; Search time 9324.54 Seconds
(without alignments)
12942.551 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttcagctgcaggctc.....ctatgacatgattacgcca 5767

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vl.*
30: em.htg.hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

```

ALIGNMENTS

RESULT	1	AX275254	Sequence 3 from Patent WO0171014.	DNA	linear	PAT 29-OCT-2001
AX275254	LOCUS	AX275254	Sequence 3 from Patent WO0171014.	5767 bp		
DEFINITION	AX275254	AX275254				
ACCESSION	AX275254	AX275254				
VERSION	AX275254.1	GI:16547674				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query
No. Score Match Length DB ID Description

1. 5767
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM034"

AX275254 Sequence
AX114873 Sequence
U89964 Cloning vec
AX191663 Sequence
AX114861 Sequence
AX191664 Sequence
AX114872 Sequence
AX114871 Sequence
U28417 Cloning vec
U02456 Cloning vec
U02437 Cloning vec
E49323 Infectious
U02454 Cloning vec
U02433 Cloning vec
U02435 Cloning vec
U02442 Cloning vec
AX195206 Sequence
U02451 Cloning vec
U02438 Cloning vec
AX244155 Sequence
AR119454 Sequence
Z32524 A.nidulans
Z32750 Expression
A82653 Sequence 3
E16636 Expression
E16677 All sequence
D84238 Nicotiana t
A93016 Sequence 4
Y07862 Cloning vec
Z32688 Expression
Z32689 Expression
Z32697 Expression
Z32699 Expression
Z32751 Expression
Z32698 Expression
Z32701 Expression
AF347016 Shuttle v
U34922 Cloning vec
L27083 Plasmid KIL
L38496 Cloning vec
AX244154 Sequence
U47294 Cloning vec
AF276982 Integrati
X65304 Cloning vec
X65305 Cloning vec

BASE COUNT					1390 a	1521 c	1495 g	1361 t						
ORIGIN														
Query Match					100.0%	Score 5767	DB 6	Length 5767						
Best Local Similarity					100.0%	Pred. No. 0								
Matches 5767					Conservative 0	Mismatches 0	Indels 0	Gaps 0						
Qy	1	agcttgcatcgctgcagg	tcaacatggtgagcagcacactcgtctactccaagaata	60										
Db	1	AGCTTGCATGCGCTGAGGTCAACATGTTGGAGCAGCACACTCTCGTCTACTCCAGAATA	60											
Qy	61	tcaagatacagctcagaagaccagagggtctattgagacttttcaacaaagggtaatat	120											
Db	61	TCAAAGATACAGTCTCAGAAGACCAGAGGGCTATTGAGACTTTTCAACAAAGGGTAATAT	120											
Qy	121	cgggaacacctcgcgattccattgcccagctatctgtcaacttcagaaagacagtag	180											
Db	121	CGGGAACCTTCTCGGATTCATTTGCCAGCTATCTGTCACTTTCAGAAAGGACAGTAG	180											
Qy	181	aaaggaagatggcttctcaaaatgccatcatgcatgataaaggaagggtatcgttcaag	240											
Db	181	AAAAGGAAGATGCTTCTACAAATGCCATCATTTGCGATAAAGGAAGGCTATCGTTCAAG	240											
Qy	241	aatgctctaccgacagtggtgccaaagatggacccccaccacagaggaaacatcgtgaa	300											
Db	241	AATGCCCTTACCGACAGTGTGTCCAAGATGGACCCGCCACCCACGAGGAACATCGTGGA	300											
Qy	301	aaagaagacgttccaaccagcttctcaagaagtgagtgatgataaacttttcaaca	360											
Db	301	AAAGAAGACGTTCCAAACCAGCTCTTCAAAGCAAGTGGATTGATGTGATAACTTTTCAACA	360											
Qy	361	aagggtaataatcgggaacacctcctcgattccattgcccagctatctgtcaacttcacga	420											
Db	361	AAGGGTAATATCGGGAACCTCTCGGATTGCCATTTGCCAGCTATCTGTCACTTCATCGA	420											
Qy	421	aaggacagtgaagaaagagatggcttctacaagtccatcatcgataaaggaaggc	480											
Db	421	AAGGACAGTAGAAGGAAGATGGCTTCTACAAATGCCATCATTTGCGATAAAGGAAGGC	480											
Qy	481	tatcgttcaagaatgcctctaccagcagtggttccaaagatggacccccaccacagagga	540											
Db	481	TATCGTTCAAGATGCTCTTACCAGCAGATGGTCCCAAGATGGACCCCAACCCACGAGGA	540											
Qy	541	acatctggaaaaagagacgtttccaaccacgcttctcaagcaagtggattgattgata	600											
Db	541	ACATCTGGAAAAAGAAAGACGTTCCAACACCGTCTTCAAAGCAAGTGAATTGATGTGATA	600											
Qy	601	tctccactgacgtaagggatgacgcacaatcccactatccttcgcaagaccccttctcta	660											
Db	601	TCTCCACTGACGTAAGGGATGACGCACAAATCCCACATATCCTTCGCAAGACGCTTCCTCTA	660											
Qy	661	tataaggaagtctatttcatttgagaggacctcgagaattaattctcaacacaacatat	720											
Db	661	TATAAGGAAGTTCAATTTTGAGAGGACCTTCGAGAAATTAATTTCTCAACACACAACATAT	720											
Qy	721	acaaacaaacgaatctcaagcaatcaagcattctacttctattgagcaatttaaatca	780											
Db	721	ACAAACAAACGAATCTCAAGCAATCAAGCATTTCTACTTCTATTGAGCAATTTAAATCA	780											
Qy	781	ttctctttaagcaaaagaatattctgaaaaatttcaaccattacgaacgtagccatg	840											
Db	781	TTTTCTTTTAAAGCAAAAGCAATTTTCTGAAAATTTTTCACCATTTTACGAACGATAGCCATG	840											
Qy	841	gctcccccgagtgctgctgacacgcttccctgagcttccccactccttctcctc	900											
Db	841	GCTCCCCCGAGTGTCTGCTGCAACGCCCTTCCCTGCTTCCCTACTCTCTCTCCTC	900											
Qy	901	ctctgctctctggtgagagtgagggtgagggcgaggagatgcagactgctggtg	960											
Db	901	CTCTGGCTCTCTGGTGGAGAGTGGGGGCTGAGGGCGGGAGGATGCAGAGCTGCTGGTG	960											
Qy	961	acggtgctggggccggctcgcggggcatttcgctgaagacccccggggccctgtctct	1020											

DB	961	ACGGTGCCTGGGGCGGGCTGCGGGGCATTTCCGCTGAAGACCCCGGGGCCCTGTCTCT	1020						
QY	1021	gcttctctgggcatcccttttgcggagccaccatggagaccccgctcgcttcttgcacacg	1080						
DB	1021	GCYTTCTGGGCATCCCTTTGGGAGCCACCCATGGGACCCCGCTGCTTCTGCCACCG	1080						
QY	1081	gagcccaagcagccttggtcaggggtggtagacgtacaaccttccagagtgctctgtac	1140						
DB	1081	GAGCCCAAGCAGCCTTGGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGTAC	1140						
QY	1141	caatatgtggaacacctataaccagagtttggagggcaccagagatgtggaacccccaccgt	1200						
DB	1141	CAATATGTGGACACCTATATACCAGGTTTTTGGGGGACCCAGAGATGTGAAACCCCAACCGT	1200						
QY	1201	gagctgagcgagagactgctgtacctcaacggtgacacacatacccccggtctacatcc	1260						
DB	1201	GAGCTGAGCGAGGACGTGCTGTACCTCAACGTGTGGACACCATACCCCGGCTACATCC	1260						
QY	1261	ccacacctgtctctgcttggaatctatgggggttggttctacagtggggacctctctctg	1320						
DB	1261	CCCACCCCTGCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCTTG	1320						
QY	1321	gagctgacagatggcgcttcttggtacagggcagagagagactgtgctggtccatgaac	1380						
DB	1321	GAGCTGTACGATGGCGCTTCTTTGGTACAGGCCGAGAGSACTGTGCTGTCTCATGAAC	1380						
QY	1381	tacgggtgggagccttggcttctcctggccctgcgggggagccgagagagcccgggc	1440						
DB	1381	TACGGGTGGGAGCCTTTGCTTCTTGGCCCTGCCGGGAGCCGAGAGGCCCGGGCAAT	1440						
QY	1441	gtgggtctctctgacagagggctggccctgagtggtgcagagaaactgagcagcttc	1500						
DB	1441	GTGGGTCTCTCTGGATCAGAGCTGGCCCTGAGTGGGTGCAGGAGAACGTGCGAGCCTC	1500						
QY	1501	gggggtgacccgcacacagtgacgctgtttggggagagcggggagccgctcgtggggc	1560						
DB	1501	GGGGGTGACCCGACATCAGTGACGCTGTTTTGGGGAGAGCGCGGAGCGCCTCGTGGGC	1560						
QY	1561	atgcacctgctgtccccgcacagccgggctgttccacagggccgtgctgcagagagcgt	1620						
DB	1561	ATGCACCTGCTGTCTCCCGCCACCGCGGGGCTGTTCCACAGGGCCGTGCTGCAGAGCGGT	1620						
QY	1621	gcccccaatggacccctgggccaagtggtggcagtgaggagggccctgcagggcaagcag	1680						
DB	1621	GCCCCCAATGGACCTCTGGGCACAGGTGGGCATGGAGAGGCGCGCTCGAGGGCACGAG	1680						
QY	1681	ctggccacacctgtgggctgctcctcagggcgacgtggtgggaatgacacagagctggt	1740						
DB	1681	CTGGCCACCTTGTGGCTGTCTCTCCAGCGGCACTGGTGGGAATGACACAGAGCTGGTA	1740						
QY	1741	gctgcttcgacacagcagcgaggtctctgtgaaccacgaatggcagctgctgcct	1800						
DB	1741	GCCTGCTTCCGACACAGCCAGCGCAGGTCTCTGTTGAACCCACCAATGGCAGCTGCTGCCT	1800						
QY	1801	caagaagcgtcttccgggttctcctctgctggtggttagatgagagacttctcagtgac	1860						
DB	1801	CAAGAAGCGCTTCTCCGGTCTCTCTGCTGCTGTGTAGATGGAGACTTCCCTCAGTGAC	1860						
QY	1861	acccagagggccctcatcaacggggagagacttccagggcctgcaggtgctgggtggtg	1920						
DB	1861	ACCCAGAGGCCCTCTCAACCGGGGAGACTTCCAGGCCCTGCAGGTGCTGGGTGGTG	1920						
QY	1921	gtgaagatgagggctcgctatttcttggttacggggcccccaggcttcagcaagaacaac	1980						
DB	1921	GTGAAGATGAGGGCTCGTATTTCTTGGTTTTACGGGCCCCAGGCTTCACAAAGACAAC	1980						
QY	1981	gagctctctacgacggggcccgagttctctggccgggtgcgggtgtccccaggta	2040						
DB	1981	GAGTCTCTCATAGCGGGGCCGAGTTCTTGGCGGGGTGCGGGTCCGGGTTCGCCAGTA	2040						
QY	2041	agtgaactggcagccgagcgtgtgctctcattacacagactggtgctcctccaggagac	2100						

Db 2041 AGTGACCTGGCAGCCGAGGCTGTGGTCTGCATTTACACAGACTGGCTGCATCCCGAGGAC 2100
 Qy 2101 ccggcagccctgaggaggccctgagcagatgtggggcgaccacaaatgtcgtgtgcccc 2160
 Db 2101 CCGGCAGCCCTGAGGGAGCCCTGAGCGATGTGGTGGCGACCAAAATGTGTGCCCC 2160
 Qy 2161 gtggccagctggctggcgactggctgcccagaggtgcccgggtacagcctaagctcttt 2220
 Db 2161 GTGGCCCAAGCTGGCTGGCGGACTGGCTGCCAGAGGTGGCCGGGTCTAGCCCTACGCTCTT 2220
 Qy 2221 gaacacagctgttccacgctctccctggccctgtggaagggtgtgcccaacggtcacag 2280
 Db 2221 GAACACGCTGTCTCCACGCTCTCTGGCCCTGTGGATGGGGTGGCCCAACGGCTACGAG 2280
 Qy 2281 atcagtgatcattttgggataccccctggacccctctgaaactacacagcgagagagaaa 2340
 Db 2281 ATCGAGTTTCATCTTTGGGATCCCCCTGACGCCCTCTCGAAAATACACGCGAGAGGAGAA 2340
 Qy 2341 atctcccccagcactgacgatactggaactggcccaactttgcccgcacagggggtcccaat 2400
 Db 2341 ATCTTGGCCCAAGCAGCTGATGCGATACTTGGGCCAACTTTGGCCGCACAGGGGATCCCCAAT 2400
 Qy 2401 gagccccgagaccacaaagcccccaaaatggccccgtacacggcggggtccagcagtac 2460
 Db 2401 GAGCCCCGAGACCCCAAGCCCCCAAAATGGCCCCGTACACGGCGGGGCTCAGCAGTAC 2460
 Qy 2461 gttagctggacctggccgctggaggtgcggggggtgcygcgccagggcctggccc 2520
 Db 2461 GTTAGCTGTGACCTGCGGCCGTGGAGTGGGGTGGCGGGGCTGCGCGCCACAGGCGCTGCGCC 2520
 Qy 2521 ttctggaaacctctcccccnaattgctcagcgtacactgataaggtaccggagctctctc 2580
 Db 2521 TTCTGGAAACCCGCTTCTCCCCAAATTTGCTCAGCGCTACTCTGATAGTACCAGGCTCTCTC 2580
 Qy 2581 aaccaatctagtagttgtctctatctatgtataaagggtatgctgatatgcactat 2640
 Db 2581 AACAACTAGCTAGATTTGCTCTATCTATATATATATTAATTAAGTATGCTGATATGCACTAT 2640
 Qy 2641 tcaaataggagcattagtagtattgttgaatgtaacttattgttatgttggaagtcacc 2700
 Db 2641 TCNAATAGGAGCATATAGGTATGTTTAAATGTCACATTTATGTTATGTTGGTAAAGTCACC 2700
 Qy 2701 taagcactccacgtacgtactgtgtgtctcttaccggctttaaataaattcttggccct 2760
 Db 2701 TAAGACATCCACGTACCTACGTTGTTGCTCTTACCGGCTTTAATAAATCTTCTGCCCC 2760
 Qy 2761 tgttccattattactaattatccctttctcactaaagaaaaattggtatcattaaagtat 2820
 Db 2761 TGTTCCATATTTACTAATATCCCTTTCTTACATAAAGAAAATTTGTTATCATTAAGTAT 2820
 Qy 2821 tagtcttagaacatataggtctttaaattgggttaggttttacaaaattaaactaatataa 2880
 Db 2821 TAGTCTTTAGAACATATAGGTCTTTAATTTGGTATGGTATGGTATGGTATGGTATGGT 2880
 Qy 2881 atgtcataaaatcccgctgggttaaacaatacagaaaatcagctcgtctattggagcga 2940
 Db 2881 ATGCTATAAATCCAGTGGTAAACAATGCAGAAAATCAGCGTCTATTTGGACCGA 2940
 Qy 2941 cagttgctattataataatggggccaccatagtagactgacaaaataaattacctgacaaca 3000
 Db 2941 CAGTTGCTATTAAT 3000
 Qy 3001 tcgtttactaaaataac 3060
 Db 3001 TCGTTTCTACTAAATAACAAACACAAAAGGGAGTGCATTTTCCAGGGCATTTTGTGAATA 3060
 Qy 3061 aaaaacagtttaaaagggtgcaatagaaaataggggtgtggaaaatagtgatttgagca 3120
 Db 3061 AAAAACAGTATAAAGGGAGTGCATATAGAAAATATAGGGGTGTGGAAATATGATTTGAGCA 3120
 Qy 3121 cgtcttgagcgaattcactggccgtctgtttacacactcgtgactgggaaaaacccctggc 3180
 Db 3121 CGTCTTGAAGCAATTCACCTGGCCGCTGTTTTACAACTGCTGACTGGGAAACCCCTGGC 3180

Qy 3181 gttaccacaacttaatgcgcttgacgacacatccccctttgcgcagctggcgtaataagcgaa 3240
 Db 3181 GTTACCACAACTTAATCGCTTGCAGACATCCCTTTCCGCAGCTGGCGTAAATAGGAA 3240
 Qy 3241 gagggccgcacccgatgcgcccctcccaacagttgcgcagcctgaatggcgaaatggcgctg 3300
 Db 3241 GAGGGCCGACCGATCGCCCTTTCCCAACAGTTGCGACGCTGAATGGCGAAATGGCGCTG 3300
 Qy 3301 atgcggtattttctcttacgcatctgtcggtattttcacacccgatatgggtgactctc 3360
 Db 3301 ATGCGGTATTTCTCTTACGGATCTGCGGTATTTACACCCGATATGTTGTCATCTC 3360
 Qy 3361 agtacaatctgtctgtatgcgcgcatagtttaagccagccccgcacacccgcacacacccgct 3420
 Db 3361 AGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCGACACCCCGCAACACCCGCT 3420
 Qy 3421 gacggccctgacgggtgtgtgtcccggtatgtgtcccggtatcgcgttacagacaagctgtgacgctc 3480
 Db 3421 GACGCGCCCTGACGGGCTTGTCTGCTCGGGCATCGCTTACAGACAAGCTGTGACCGTC 3480
 Qy 3481 tccgggagctgcatgttcagaggtttttcacccgtcatcacccgaaacgcgcgagagaaag 3540
 Db 3481 TCCGGGAGCTGCATGTGTACAGAGTTCACCGTCTATCACCGAAAACGCCGAGACGAAAG 3540
 Qy 3541 ggcctgtgatacgctctattttataggttaatgtcatgataataatgggttctctagcgt 3600
 Db 3541 GGCCCTGCTGATACGCTATTTTATAGGTAAATGTATGATTAATATGTTTCTTAGACG 3600
 Qy 3601 tcaagtgccacttttcggggaaatgtgcgggaaacccctatttcttatttttctaaata 3660
 Db 3601 TCAGGTGCACCTTTTCGGGAAATGTGCGGAAACCCCTATTTGTTATTTTCTAAATA 3660
 Qy 3661 cattcaaatatgtatccgctcatgagacaataaccctgataaactctcaataatgta 3720
 Db 3661 CATTCAAATATGTATCCGCTCATGAGACAAATAACCCCTGATAAATGCTTCAATAATGTA 3720
 Qy 3721 aaaaggaagtagtatgagtattcaacatttcogtgcgccttattcccttttttgcggca 3780
 Db 3721 AAAAGGAAGATGATGATTTCAACATTTCCGCTGTCGCGCTTATTCCTCTTTTTCGGCA 3780
 Qy 3781 ttttgcctctctgttttgcctcaccagaaacgcgtgggaaagtaaaagatcgtgaagt 3840
 Db 3781 TTTTCCCTTCTCTGTTTGTCTCACCAGAAACGCTGGTGAAGATAAAAGATGCTGAAGAT 3840
 Qy 3841 cagttgggtgacgagtggtttacatcgaaactggaactcctcaacagcggtaagatccttgag 3900
 Db 3841 CAGTTGGGTGCACGAGTGGTTTACATCGAATCGAATCGATCTCAACAGCGGTAAAGATCCTTGAG 3900
 Qy 3901 agtttcgccccgaagaacggttttccaaatgtagagcaacttttaaagtttctgctagtggc 3960
 Db 3901 AGTTTCCGCCCGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCG 3960
 Qy 3961 gcggtattatccggtattgacggcggaagcgaactcggctgcgcgcatacactattct 4020
 Db 3961 GCGGTATTATCCCGTATTGACCGCCGGGAAGAGCAACTCGGTCGCGCGCATACACTATTCT 4020
 Qy 4021 cagaatgactgtgtgactcaccagtcacagaaaagcactcttacggatggcatgaca 4080
 Db 4021 CAGAAATGACTGGTTGAGTACTCACCGATCACAGAAAAGCATCTTACGGATGGCATGACA 4080
 Qy 4081 gtaagaagaattatgcagtgctgccataaccaatgagtatacactgcggcccaactactt 4140
 Db 4081 GTAAGAAGAAATATGCACTGCTGCCATTAACCATGAGTGAACACTGCGGCCCAACTTACTT 4140
 Qy 4141 ctgacaacagatcggaggccgaagagctaaacgcttttttgcacaaactgggggatact 4200
 Db 4141 CTCACAACGATCGGAGGACCCGAAGAGCTTAACCGCTTTTTCACAACTGSGGGGATCAT 4200
 Qy 4201 gtaactcgccttgatcgttgggaacggagctgaaatgaagccatataccaaacgacgagct 4260
 Db 4201 GTAACCTCGCTTGCATCTGTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGCGT 4260

```
QY 4261 gacacacgatgcttagcaaatggaacaacacgttgccgcaaaactattaaactggcgaaacta 4320
Db 4261 GACACCACGATGCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTA 4320
QY 4321 ctactactagctcccgcaacaataatagactggatggagcggaataaagtgcaggga 4380
Db 4321 CTACTCTAGCTTCCCGCGCAACAATAATAGACTGGATGGAGCGGGATAAAGTTGCAGGA 4380
QY 4381 ccactctgctgcgcctccgctccgctggctggtttattctgataaaactggagccggt 4440
Db 4381 CCACTTCTGCGCTCGGCCCTTCGGCTGGCTGGTGTATTGCTGATAAATCTGGAGCCGGT 4440
QY 4441 gaggctgggtctcggtgatactgagcactggggccagatggtgaagccctcccgatc 4500
Db 4441 GAGCGTGGGTCTCGCGGTATCATTTGCAGCACCTGGGGCCAGATGGTAAGCCCTCCCGTATC 4500
QY 4501 gtgattatcacagcagcgaggatcaggcaactatgatgaacgaataagacagatcgct 4560
Db 4501 GTAGTTATCTACAGCAGCGGGAGTACGCCAATCTATGGATGAACGAATAAGACAGATCGCT 4560
QY 4561 gagatgggtcctcactgattaagcattggttaactgtcagaccgaagtctactcatata 4620
Db 4561 GAGATAGGTGCTCAGTATTAAGCATTTGCTAAGTGTGACAGCAAGTTTACTCATATATA 4620
QY 4621 ctttagattgattaaactctattttaaataaggtatagtggaagatccctttt 4680
Db 4621 CTTTAGATTGATTAAACTTTCATTTTAAATTTAAAGSATCTAGGTGAAGATCCCTTTT 4680
QY 4681 gataatctatgacccaaaatcccttaacgtgagtttccgtccactgagcgtcagacccc 4740
Db 4681 GATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCC 4740
QY 4741 gtgaaaaagatcaaaagatctcttgatgataccttttttctgpcgctaactctgcttg 4800
Db 4741 GTAGAAAAGATCAAAAGATCTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTG 4800
QY 4801 caaacaacacacacacacacacacacacacacacacacacacacacacacacacacac 4860
Db 4801 CAAACAAAACACACACACACACACACACACACACACACACACACACACACACACACAC 4860
QY 4861 cttttccgaaggttaactggctcagcagcagcagcagcagcagcagcagcagcagcagc 4920
Db 4861 CTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCCAGACATACCAATACTGCTTCTAGTG 4920
QY 4921 tagccgtagtttagccaccacttcaagaactctgtagcaccgctacatacctcgtctg 4980
Db 4921 TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTTACATACCTCGCTCTG 4980
QY 4981 ctaactcgtttaccagtgctgctgcccagtgccagtgccagtgccagtgccagtgccagtg 5040
Db 4981 CTAATCCTGTTTACCAGTGGCTGCTGCGAGTGGCGATAAGTGTCTTACCGGGTTGGAC 5040
QY 5041 tcaagacgatagttaccgggataaggcagcagcagcagcagcagcagcagcagcagcagc 5100
Db 5041 TCAAGACGATAGTTACCGGATAGGCGCAGCGCTCGGGCTGAACGGGGGTTGCTGCACA 5100
QY 5101 cagcccgcttgagcgaacacactacacccgaactgagatcacctacagcgtgagctatga 5160
Db 5101 CAGCCCGCTTGAGCGCAACGACCTACACCGAACTGAGATACCTACAGCGCTGAGCTATGA 5160
QY 5161 gaaagcgcacgcttcccggaaggagaaaggcagcagcagcagcagcagcagcagcagcagc 5220
Db 5161 GAAAGGCCACGCTTCCGGAAGGAGAAAGGCGGACAGGTATCCGGTAAGCGCAGGGTC 5220
QY 5221 ggaacaggagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5280
Db 5221 GGAACAGGAGAGCGCAGCGAGCTTCCAGGGGAAACCGCTGGTATCTTTATAGTCTCT 5280
QY 5281 gtcggggtttccacacctctgacttgagcgtcgatatttggatgctgctgagggggcg 5340
Db 5281 GTCGGGTTTCCGACACCTCTGACTTGAGCGTCTGATTTTGTGTGCTCGTCTGAGGGGCGG 5340
QY 5341 agcctatggaaaaacccagcaacgcggccttttttaccagggttctggtgacctctcagta 5400
```

```
Db 5341 AGCCTATGGAAAACGCCAGACCGCGCTTTTACGGTCTCTGGCCTTTTGGTGCCT 5400
QY 5401 ttgtctcacatgttcttctctggttattccctgattctggtataaccgtattaccgccc 5460
Db 5401 TTTGCTCACATGTTCTTTCTCTGGTATTCCCTGATTCTGTGGATAACCGTATTACCGCC 5460
QY 5461 tttagtgagctgataccgctcgcgcagcagcagcagcagcagcagcagcagcagcagcagc 5520
Db 5461 TTTGAGTGAGCTGATACCGCTCGCCGACGCGAAGACCGCAGCGCAGCGAGTCACTGAGC 5520
QY 5521 gaggaaagggaagagcggcccaataatcgcaaacccgctctcccgcgcttggcgagattcat 5580
Db 5521 GAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCTTGGCGCATTTTAT 5580
QY 5581 taatgcagctggcagcagcagcttcccgactggaagcggcagcagcagcagcagcagcagc 5640
Db 5581 TAATGACAGCTGGCAGCAGCAGCTTCCGACTTGGAAACCGGCGCAGTGAAGCGCAACGCA 5640
QY 5641 aatgtgagttagctcactcatttaggcaccccgctttacacctttatgcttccggctcgt 5700
Db 5641 AATGTGAGTTAGCTCACTCATTTAGGCACCCAGCGCTTTACACTTTATGCTTCCGGCTCGT 5700
QY 5701 atgttgtgtgaattgagcgggataaacaatttcacacaggaacacagctatgacctgat 5760
Db 5701 ATGTTGTGTGGAATTGTGAGCGGATAACAATTTTCACACAGGAAACAGCTATGACCATGAT 5760
QY 5761 tacqcca 5767
Db 5761 TAGCCCA 5767

RESULT 2
AX114873
LOCUS AX114873 3858 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 33 from Patent WO0129208.
ACCESSION AX114873
VERSION AX114873.1 GI:14031815
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 3858)
AUTHORS Kuehn, R., von Melchner, H. and Altschmied, J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 33 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE); Frankgen Biotechnologie AG (DE)
FEATURES
source
location/Qualifiers
1..3858
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="vector PRK50"
BASE COUNT 943 a 959 c 994 g 962 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaactgcccgtctgttttacaacgctgctgactggggaacccctggcgttaccacaac 3190
Db 1187 CGAATTCACTGGCGCTGCTTTTACAACGCTGACTGGGAAAACCCCTGGCGTTACCCAAC 1246
QY 3191 ttaatgccttcagcacatcccccttccagctggcgtaataagcgaagagcccgca 3250
Db 1247 TTAATGCGCTTGCAGCACATCCCGCTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCGCA 1306
QY 3251 ccgacgccccttcccaacagttgcgagcctgaatggcgaatggcgtgatgctggtatt 3310
Db 1307 CCGATCGCCCTTCCCAACAGTTGCGCAGCGCTGAATGCGAATGGCGCTGATGCGGTATT 1366
QY 3311 ttctcctaagcgtctgctggttatttcacacggcataatggtgcactctcagtaaatct 3370
```

2638/38

Db 1367 TTTCTCTTACGCATCTGTGCGGTATTTTCACACGCATATGTGCTGCTCTAGTACATCT 1426
Qy 3371 gctctgtagcgcgcatagttaagcagccccgacaccccgcccaaccccgctgacgcgcct 3430
Db 1427 GCTCTGATGCGCATAGTTAAGCAGCGCCGACACCCGCCAACACCCGCTGACGCGCCT 1486
Qy 3431 gacgggctgtctctcccgcatccgcttacagacaagctgtgacctctccgggagct 3490
Db 1487 GACGGGCTGTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGAGCT 1546
Qy 3491 gcatgttcagaggttttcacgcgtcatcacccgaaacgcgcagacgaagggcctcgtga 3550
Db 1547 GCATGTGTACAGGTTTTCACCGTCATCACCGAAGCGCGAGAGAAAGGGCCTCGTGA 1606
Qy 3551 taccgctattttatagtgtaagtcatgataataatgtttcttagcgtcagtggtgca 3610
Db 1607 TACGCCATATTTTATAGGTTAATGTGATGATGAATAATGGTTCTTAGACGTCAGGTGCA 1666
Qy 3611 cttttccgggaaatgtgcgcggaacccctatttggttatttttctaaatacatcaata 3670
Db 1667 CTTTTCCGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTCTTAAATACATTTCAATA 1726
Qy 3671 tgcacgcctcatgagacaaataacccctgataaagtctcaataatgaaaaaggaaga 3730
Db 1727 TGTATCCGCTCATGAGACATTAACCTCATTAATGCTTCAATATATTTGAAAAGGAAGA 1786
Qy 3731 gtatgagttatcaacatttcgctgcgccttattcccttttttgcggcattttgcccctc 3790
Db 1787 GTATGAGTATTCACATTTCCGTGTGCGCCCTATTTCCTCTTTTTCGCGCATTTGCCCC 1846
Qy 3791 ctgttttgcctcaccggaacacgctggtgaaagttaaaagatacgtgaagatcaagttgggtg 3850
Db 1847 CTGTTTTTGCCTCACCCAGAAACGCTGTGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 1906
Qy 3851 cagcagtggtgtacatcgaaatgcatccaacagcgttaagatcctgagagtttccgc 3910
Db 1907 CACGATGGGTACATCGAATGATCTCAACAGCGGTAAAGTCTTGAGAGTTTTCGCC 1966
Qy 3911 ccgaagacggttttccaatgatgagcaacttttaagttctgtatgtggtgcggtattat 3970
Db 1967 CCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGTATGTGCGCGGTATTAT 2026
Qy 3971 cccgtattgacgcggggaagagaacacgctgcgcgcatacacatattctcagaatgact 4030
Db 2027 CCGGTATTGACGCGCGGGAAGAGCAACTCGGTGCGCGCATACATATTCTCAGAAATGACT 2086
Qy 4031 tggttgagttactcaccagttcacagaaaagcatcttacggatggcatgacagttaagagaat 4090
Db 2087 TGGTTGAGTACTACCACTCAGANNAAGCATCTTACGGATGGCATGACATGAGAGAT 2146
Qy 4091 tatgcagttgctgccaataaccatgagtgataacactgcggccaacttactcttgacaacga 4150
Db 2147 TATGACAGTGTGCCATACCATAGTATACACTGCGGCCAACTTACTTCTGACAACGA 2206
Qy 4151 tcggaggaccgaagagtaaccgctttttgcacaacatgggggatacatgtaactcgcc 4210
Db 2207 TCGGAGGACCGAAGAGGTAAACCGCTTTTTTTCACAACATGGGGGATCATGTAACTCGCC 2266
Qy 4211 ttgatcgttgggaacggagctgaatgaagccataccaaacgacgagcgtgaacacacga 4270
Db 2267 TTGATGTTGGGACCGGAGCTGAATGAAGCCATACCAACACGAGCGTGACACCAAGA 2326
Qy 4271 tgcctgtagcaattggcaacaacggttgcgcaaacatttaactggcggaacttacttacttag 4330
Db 2327 TGCCCTGTAGCAATGGCAACAGTTTGCAGAACTATTAACTGGCGAACTACTTACTCTAG 2386
Qy 4331 cttcccgcaacaataatagactggaatggaggcgaataaagttgcaggaaccactctcgc 4390
Db 2387 CTTCGCCGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTTCAGAGGACCACTTCGC 2446
Qy 4391 gctcgcccttccgctgagctgggtttattgtctgataaatctggagccggtgagcgtgggt 4450
|||||

Db 2447 GCTCGGCCCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGGTGAGCGGTGGGT 2506
Qy 4451 ctgcgggtatcatctgacgactggggccagatggtaagccctcccgatatcgtagttatct 4510
Db 2507 CTGCGGTATCATTTGACGACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 2566
Qy 4511 acacgacgggagtcaggcaactatggatgaacgaaaaagacagatcgtcgagatagggtg 4570
Db 2567 ACACGACGGGAGTCAGGCAACTATGGGATGAACGAAATAGACAGATCGCTGAGTAGGTG 2626
Qy 4571 cctcactgatlaagcattggttaactgctcagaccagtttactcatataacttttagattg 4630
Db 2627 CCTCACTGATTAAGCATTTGTAACCTGTCAGACCAAGTTTACTCATATATATCTTTAGATTG 2686
Qy 4631 atttaaaacttcatttttaatttaaaagatctaggtgaagatccctttttgataatctca 4690
Db 2687 ATTTAAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATATCTCA 2746
Qy 4691 tgacaaaaatcccttaacgtagtttcttcactgagcgtcagaccccgtagaaaaaga 4750
Db 2747 TGACCAAAATCCCTTAACGTGAGTTTTCCTTCCACTGAGCGGTACAGCCCGTAGAAAAGA 2806
Qy 4751 tcaaggagctctcttgagatcccttttctgcgcgttaactgctgctgctgcaaaacaaa 4810
Db 2807 TCAAGGATCTCTTGAGATCCTTTTTCGCGCGTAACTGCTGCTGCTGCTGCAACAAAA 2866
Qy 4811 aacacacgctaccagcgtggtttgttcgcgatcaagagctaccaactctttttccga 4870
Db 2867 AACCACCGCTACCAGCGGTGTTGTTGCGCGATCAAGAGCTACCAACTCTTTTTCGA 2926
Qy 4871 aggttaactggcttcagcagcagcagataacaaaaactgctcttcttagtgagcgtagt 4930
Db 2927 AGGTAACTGGTTCTACGACGAGCGAGATACCAAAATACTGTCTTCTAGTGTAGCGGTAGT 2986
Qy 4931 taggcacacacttcaagaaactctgtagcaccgctcacatcactcctctgcttaactcgt 4990
Db 2987 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTGCTGTAACTCTGT 3046
Qy 4991 taccagttggctgctgccagtgcggaataagtcgtcttaccgggttggaactcaaacagat 5050
Db 3047 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGTTGGACTCAAGACGAT 3106
Qy 5051 agttaccggataagcgcagcgttcgggctgaacgggggttcggtgcacacagccagct 5110
Db 3107 AGTTACCAGATAAGCGCAGCGGTGCGGTGAACCGGGGTTTCGTGCACACACGCCAGCT 3166
Qy 5111 tggagcgaacacctcacccgaactgagatacctacagcgtgagctagagaaagcga 5170
Db 3167 TGGACGCAACGACCTACACCGAATGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 3226
Qy 5171 cgcttccggaagggaagcgggacaggtatccggtaagcgcggggtcggaacagagag 5230
Db 3227 CGCTTCCGAAGGGAGAAAGGGGACAGGTATCCGGTAAGCGCAGGTTCGGAACAGGAG 3286
Qy 5231 agcgaacagggagcttcagggggaacgcctggtatcttcttagtctcgtcgagtttc 5290
Db 3287 AGCGACAGGAGAGCTTCCAGGGGAAACGCTGCTATCTTTATAGTCTCTGCGGTTTC 3346
Qy 5291 gccactctgactgagcgtgatttttctgctgctcagggggcgagcgttatgga 5350
Db 3347 GCCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGA 3406
Qy 5351 aaaaacgcagcaacgcggcctttttacggttctcgtgctgcttggccttttgcctcaca 5410
Db 3407 AAAACGCCAGCAACCGCGCTTTTACGGTTCCTGCGCTTTTGTGCGCTTTTGTCTACA 3466
Qy 5411 tgttttctcgtgcttatccccctgattctgtggaataaccgtatttaccgcttttgagtag 5470
Db 3467 TGTCTTTCTCTGCTATATCCCTGATTCTGTGGATAAACCGTATTACCGCTTTTGAGTAG 3526
Qy 5471 ctgatccgctcgcgcagcgaacgacgagcgcagcgtcagtcagtcagcggagagcgg 5530
Db 3527 CTGATACCGCTCCCGCACCCCAACGACCGGACGCGAGTCACTGACGAGGAGACCGG 3586
|||||

```
QY 5531 aagagcgcccaataagcaaacccgctctccccgcgcgcgttgccgattcattaatgcagct 5590
|||||
Db 3587 AAGAGCGCCCAATACGAAACCGCCTCTCCCGCGGCTTGCGCGATTCAATATGCAGCT 3646
|||||
QY 5591 ggcagacaggtttcccgactgaaagcggcgagtgagcgcaacgaataatgagtt 5650
|||||
Db 3647 GGCACGACAGGTTCCCGACTGAAAGCGGCGAGTGAGCGCAACGAATTAATGTGAGTT 3706
|||||
QY 5651 agtcactcattagcaccacccagctttacacttatgtcttcgcgcgtcgtatgtgtg 5710
|||||
Db 3707 AGCTACTACTTAGGACACCCAGGCTTACACTTATGCTTCGGGCTCGTATGTGTGTG 3766
|||||
QY 5711 gaattgtgagcggaataacatttcacacaggaacagctatgaccatgattacgcca 5767
|||||
Db 3767 GAATTGTGAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTAGGCCA 3823
|||||
```

RESULT 3

```
UC089964
LOCUS          UC089964          4473 bp      DNA      circular SYN 03-SEP-1997
DEFINITION    Cloning vector pG5CAT, complete sequence.
ACCESSION     U89964
VERSION       U89964.1 GI:2293130
KEYWORDS
SOURCE        unidentified cloning vector.
ORGANISM      Holtz, A. and Lou, Y.
REFERENCE     1 (bases 1 to 4473)
AUTHORS      Holtz, A. and Lou, Y.
JOURNAL       PG5CAT complete sequence
TITLE        Unpublished
REFERENCE     2 (bases 1 to 4473)
AUTHORS      Holtz, A. and Lou, Y.
JOURNAL       Direct Submission
COMMENT       Submitted (19-FEB-1997) CLONTECH Laboratories, Inc., 1020 East
Meadow Circle, Palo Alto, CA 94303-4230, USA
This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303-4230, USA. To place an
order call (415) 424-8222 or (800) 662-2566, extension 1.
International customers, please contact your local distributor. For
technical information, call (415) 424-8222 or (800) 662-2566,
extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. If you suspect there is an
error in this sequence, please contact CLONTECH's Technical Support
Department at (415) 424-8222 or (800) 662-2566, extension 3 or
E-mail TECH@CLONTECH.COM
```

FEATURES

```
source      Location/Qualifiers
1..4473
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
gene        120..779
           /gene="CAT"
CDS         120..779
           /gene="CAT"
           /codon_start=1
           /product="chloramphenicol acetyltransferase"
           /protein_id="AAB68658.1"
           /db_xref="GI:2293131"
translation="MEKKITGYTTVDISQWHRKEHFEAFQSAQCYNTQVOLDITAF
LTKVKNKHVPYPAFIHLARLMAHPEFRMAKDELVIDWSVHPCYTVFHEQPTF
SSLWSEYHDDRFQFHHYSODVACVGENLAYFPKGFIEFMFVSANPVSTSEDLNV
ANMDFNFAVPVFMGYKYYTGDKVLMPLAIQVHHAVCDGFHVRMLNELQQYCDENQGG
A"
gene        2337..3197
           /gene="ampR"
CDS         2337..3197
           /gene="ampR"
           /function="ampicillin resistance"
           /codon_start=1
```

```
/transl_table=11
/protein_id="AAB68658.1"
/db_xref="GI:2293132"
/translation="MSIOHFRVALIPFAAFLCPVFAHPETLVKVKDAEDOLGARVGY
IELDLNSGKILESFRPEFRPMNFKVLICGAVLSRIDAGOEOLGRIHYSONDLVE
YSPVTEKHLTDGMYVRELCASAIWMSDNTAANLLTTTGGPKELTAFLHNMGDHTVRL
DRWPELNEAIPNDRDITMPVAMATTIKLLTGLLELLTSLRQQLIDHMEADKVAGPL
LRSLALPAWFIADKSGAGERSGIIAALGPDGKPSRVIVITTSQATMDERNQRIA
EIGASLIRHW"
BASE COUNT    1183 a    1044 c    1046 g    1200 t
ORIGIN
```

```
Query Match      45.7%; Score 2637; DB 12; Length 4473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactgcccgtcgcttttacaacgctgctgactgggaaaaacccctggcggttaccacaac 3190
|||||
Db 1735 CGAATTCACCTGCGCGTCTGTTTACAACTGCTGACTGGAACACCCCTGGCGTTTACCACAAC 1794
|||||
QY 3191 ttaategccttgccagcacacatccccctttccagagctggcgtaataagcgaagagcccgca 3250
|||||
Db 1795 TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATACGGAAGAGCCCGCA 1854
|||||
QY 3251 ccgatacgcccttcccaacagttgctgcagcctgaatggcggaatggcgccctgatgcggtatt 3310
|||||
Db 1855 CCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGCGAATGGCGCTGATGCGGTATT 1914
|||||
QY 3311 ttctcctacgactctgtcggtattttcacacgcgcatacggcgactctcagtaacaatct 3370
|||||
Db 1915 TTCTCCTTACGCATCTGTCGGTATTTCACACCGCATATGGTGCACTCTCAGTACATACT 1974
|||||
QY 3371 gctctgatccgcatagttaagccagcccgacaccccgcaacaccccgctgacgcgcct 3430
|||||
Db 1975 GCTCTGATCCCGCATAGTTAAGCCAGCCCGCACACCCCGCAACACCCCGTACGCCCCCT 2034
|||||
QY 3431 gacgggctgtctgctccggcgcctccgtttacagacaagctgtgaccctctccggagct 3490
|||||
Db 2035 GACGGGCTTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 2094
|||||
QY 3491 gcattgtcagagggttttcacgcgtcatccgaaacgcgcgagacgaaagggcctcgtga 3550
|||||
Db 2095 GCAATGTCAGAGGTTTTCACCGTCAACACCGCAACGCGAGACGAAAGGGGCTCGTGA 2154
|||||
QY 3551 tacgcctattttataggttaatgctatgataataatggtttctttagacgtcagtgagca 3610
|||||
Db 2155 TAGCCCTATTTTATAGGTAAATGTCATGATATAATGTTTCTTAGAGCTCAGGTGGCA 2214
|||||
QY 3611 ctttcggggaaatgtgcgcggaacccctatttggttatttttcttaataacattccaata 3670
|||||
Db 2215 CTTTTCGGGAAATGTGCGCGGAACCCCTATTGTTTATTATTCTTAATACATTCAATA 2274
|||||
QY 3671 tgtatccgctcagagacaataaccctgataaaatggttcaataataattgaaaaagaaga 3730
|||||
Db 2275 TGTATCCGCTCATGAGACAATAACCCCTGATAATGCTTCAATAATAATTGAAAAGGAAGA 2334
|||||
QY 3731 gtagagattcaacaatttcgctgcgccttattcccttttttgcggcatcttgccttc 3790
|||||
Db 2335 GTATGAGTATTCAACATTTCCGCTGTCGCCCTTATTCCCTTTTTCGGGCATTTTGGCTTC 2394
|||||
QY 3791 ctgtttttctcaccacagaaacgctgggtgaaagtaaaagatcgtgaagatcagttgggtg 3850
|||||
Db 2395 CTGTTTTTCTCACCACAGAACGCTGGTGAAGTAAGATGCTGAAGATCAGTTGGGTG 2454
|||||
QY 3851 cscgagtggtttacatcgaactggatctcaacagcggtaagaatcccttgagaggttttcgcc 3910
|||||
Db 2455 CACGAGTGGGTTCATATCGAACTCGAATCTCAACAGCGGTGAAGATCCCTTGAGAGTTTTTCGCC 2514
|||||
QY 3911 ccgaagaacgcttttccaatgatgcacacttttaaaagtctcgtatgtgcgcggtattat 3970
|||||
Db 2515 CGGAAGAAGCTTTTCCAAATGATGAGGACACTTTTAAAGCTTCTGCTATGTGCGCGGTATTAT 2574
|||||
```

Qy	3971	cccgata	tcgacgcg	cgccggaagagcaactcgg	tcgpcgcatacactattctcagaatgact	4030
Db	2575	CCCCATATTGACGCGCGGCAAGACGAACCTCGGTGCGCGCATACACTATTCTCAGAATGACT	2634			
Qy	4031	tgggttagtactaccagttcacagaaaagcattctacgga	tggccatgacagtaagagaat	4090		
Db	2635	TGGTGTAGTACTCACCAGTGCACAGAAAAGCATCTTACCGATGGCATGACATAGAGAAT	2694			
Qy	4091	tatcgagtctgcataaaccatagagtga	taaacactgcgcgccaacttacttctgcacaacga	4150		
Db	2695	TATGACGTGCTGCCATAACCATGAGTGATAACACTTGGCGCAACTTACTTCTGCACACGA	2754			
Qy	4151	tcggagagaccgaaggagctaaacgcgttttttgcacaa	catgtaactgcgccttaactgcgc	4210		
Db	2755	TCGGAGACGCAAGGAGACTTAACCGCTTTTTCACAA	CTGGGGATCATGTAATCTGCC	2814		
Qy	4211	ttgaatgttgggaaccggagctgaatgaagccat	accacaaacgacgagcgtagacacacga	4270		
Db	2815	TTGATCGTTGGGAACCCGGAGCTGAATGAAGCCAT	TACCAACGACGAGCGTGACACCAACGA	2874		
Qy	4271	tgcctgtagcaaatggcaacaacgcttcgcacaa	actattaaactggcgaaacttacttacttactctag	4330		
Db	2875	TGCGTGTAGCAATGGCAACAACGTTGCGCAAACT	ATTAACTGGCGCAACTTACTTACTCTAG	2934		
Qy	4331	cttcocggcaacaattaaatagactggatggagcg	cggtataaagttgcagggaccactcttcgc	4390		
Db	2935	CTTCCCGGCAACAATTAATAGACTGATGGAGCGCG	ATAAAGTTGCAGGACCACTTCTGCG	2994		
Qy	4391	gctcgcccttcgcgtgctggtttattgctgata	ataatctggagcgggtgagcgtagggt	4450		
Db	2995	GCTCGGCCCTTCGCGCTGGCTGTGTTTATTGCT	TGATTAATCTGGAGCGCGTGAGCGTGCGGT	3054		
Qy	4451	ctcgcggtatcattgtagcactggcgccagat	tggtaaagccctccgctatcgtagttatct	4510		
Db	3055	CTCGCGGTATCAATTGCAGCACTGGGCGCCAG	TGGTAAGCCCTCCGCTATCTGATTAATCT	3114		
Qy	4511	acacgcggggagtcagggcnaactatgga	tgaacgaaaa	tagacagatcgctgagataggtg	4570	
Db	3115	ACAGCAGCGGGAGTCAAGCAACTATGGATGA	ACAGAAATAGACAGATCGCTGAGATAGGTG	3174		
Qy	4571	cctcactgataagacttagtgaactgcagac	ccaagtttactcatatactttagattg	4630		
Db	3175	CCTCAGTGTATAGCAATGGTAAGTCTCAG	ACCAAGTTTACTCATATATACTTTAGATG	3234		
Qy	4631	atttaaaactctttttaaattaaaggatc	tagttagaagatcctttttgataatctca	4690		
Db	3235	ATTTAAACCTTCATTTTAA	TATAAAGGATCTAGTGCAAGATCCTTTTGTGATTAATCTCA	3294		
Qy	4691	tgacaaaaa	tcctctaacgttgagttt	tcctcactgagcgcagacccogtagaaaaga	4750	
Db	3295	TGACCAAAATCCCTTAACGCTGAGTTTTCG	TTCCTTCCACTGAGCGTCAGACCCCGTAGAAAAGA	3354		
Qy	4751	tcaaggatctcttgagatccttttttctgcg	gtaactctgctgcttgcaacaaaaa	4810		
Db	3355	TCAAAGGATCTCTTGAGATCCCTTTTTC	TGCGCGGTAAATCTGTGCTGTGCAACCAAAA	3414		
Qy	4811	aaccacgcgtaccagcggtgggtt	gttcgccgga	tcaagactcaacacttttttcoga	4870	
Db	3415	AACCACCGCTACCAGCGGTGGTTGTT	TGTCGCGGATCAAGACTACCAACTCTTTTTCOGA	3474		
Qy	4871	aggtaactggcttcagcagagcgcagata	taccaaaa	tactgtccttcttagtgtagcgttagt	4930	
Db	3475	AGGTAAGTGCCTTCAGCAGAGCGCAGAT	ACCAAAATAC	TGTCTTCTAGTGTAGCCGTAGT	3534	
Qy	4931	taggccaacacttcaagaactctgtagac	cgctcaatacctcgctctgctaactctgt	4990		
Db	3535	TAGGCCACCACTTCAAGAACTCTGT	TAGCACCGCCCTACATACTCGCTCTGCTAATCTCTGT	3594		
Qy	4991	taccagtggctgctgcagtgggcgata	gaagtcgtcttaccggggttggaactcaagacgat	5050		
Db	3595	TACCAGTGGCTGTCGACGATGGCGGAT	AAGTCGTCTCTTACC	CGGGTGGGACHTCAAGACGAT	3654	
Qy	5051	agttaccggataaggcgacgcggtgcgg	ctgaacgggggggttcgtgcacacagccccagct	5110		

Db	3655	AGTTACCGGATAAGGGCGACGGCTGGGGCTGAACGGGGGGTTCTGTGCACACAGCCAGCT	3714
Qy	5111	tggagcgaacgacctcacgcgaactgagatacctcacgcgtgagctatgagaaagcgcca	5170
Db	3715	TGGAGCGAGACGACTACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA	3774
Qy	5171	cgcttccggaagggagaaagcgagaggtatcccgtaagcggcgaggtcggaacaggag	5230
Db	3775	CGCTTCCGAAGGAGAAAGCGCGACAGCTATCCGGTAAGCGGACGGCTCGGAACAGGAG	3834
Qy	5231	agcgacagggaggttccagggggaaacgcgctggtatctttatagtctcggggttc	5290
Db	3835	AGCGACGAGGGAGCTTCCAGGGGGAACGCCTGTGTATCTTTATAGTCTCTCGGGTTTC	3894
Qy	5291	gceacctgacttgagctcgagtttttggatgctctcaggggggcgagacctatgga	5350
Db	3895	GCCACCTCTGACTTGAGCTCGATTTTTGTGTATCTCTCAGGGGGCGAGCCCTATGGA	3954
Qy	5351	aaacgcgcagcaacgcggcccttttacggttctcctggccttttggccttttgcctcac	5410
Db	3955	AAAACGCCAGCAACGGCGGCTTTTTACGGTCTCTGGSCCTTTTGCTGGCCTTTTGC	4014
Qy	5411	tgtctttctcggttatccccctgattcttggtgataaacgctattaccgcctttgagtgag	5470
Db	4015	TGTTCTTTCTGGGTTATCCCCCTGATTCTGTGGATAACCGTATTATACGCCCTTTGAG	4074
Qy	5471	ctgatacgcctcgccgcagccgaacgcagcgacgcgagtcagtcagcgaggaagcgg	5530
Db	4075	CTGATACCGCTCGCGCAGCGCAACGACGCGCAGCGATCTCAGTGAGCGAGGNAACGG	4134
Qy	5531	aagagcgcccaatcacgaaacgcctctccccgcgcgttggcgcgattcattaatgcagct	5590
Db	4135	AAGAGCGCCCAATACGAAACCCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCT	4194
Qy	5591	ggaacgacaggtttccgcgactgaaagcgggcagtgagcgcaacgcgaattaatgtgagtt	5650
Db	4195	GGCACGACAGGTTCCCGACTGGAAAGCGGCAGTGAGCGCAACGCAATTAATGTGAGTT	4254
Qy	5651	agctcactcataggcaccgcaggtttacactttatgcttccggctcgatgttggtg	5710
Db	4255	AGCTCACTCATAGGCACCCAGGCTTTACACTTATGCTCCCGCTCGTATGTTGTGTG	4314
Qy	5711	gaattgtgacggataaacatttcacagggaaacagctatgacatgattacgcca	5767
Db	4315	GAATTGTGACGGGATACAAATTTACACAGGAAACAGCTATGACCATGATTACGCCA	4371

RESULTS

Accession #	LOCUS	DEFINITION	Sequence	Size	Source	Accession #
AX191663	LOCUS	AX191663	Sequence	4960 bp	DNA	PAT 15-AUG-2001
			from Patent	WO0149832.	linear	

ACCESSION	AX191663
VERSION	AX191663.1
GI:	15209844

KEYWORDS
SOURCE
synthetic construct.

ORGANISM synthetic construct
artificial sequence

REFERENCE
AUTHORS
1 (bases 1 to 4960)
Schwientek P

TITLE Transduction of recombinases for inducible gene targeting

ARTEMIS Pharmaceuticals GmbH (DE)
Patent: WO 0149632 A 29.12.00-2001;
Locations/Qualifications

FEATURES	Location/Qualifiers
source	1, .4960

```
/organism="synthetic construct"  
/db_xref="taxon:32630"
```

BASE COUNT	1225 a	1213 c	1296 g	1226 t
/note=vector pCMV-I-Cre-pA"				

ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 4960;

Qy	4391	gctcgcccttc	cggtcgctgg	ttatttgctg	tataatct	gagccgg	tgagcgtgggt	4450
Db	4300	GCTCGGCCCT	TCCGGCTGG	CTGTTATT	TGCTGATAAA	CTCGAGCCGG	TGAGCTGGGT	4359
Qy	4451	ctcgcggtat	catctgcag	caactggg	ggccagat	ggtaagc	ctcccgatatcgt	4510
Db	4360	CTCGCGGTAT	CTATGAGAC	CTGGGGCAG	ATGGTAA	GCCCTCCG	TATCGTAGTTATCT	4419
Qy	4511	acacgacggg	aggtcagg	caacttgg	atgaac	gaataag	acagatcgc	4570
Db	4420	ACACGACGGG	AGTACAGCA	CTATGGAT	GAACGAAT	AGACAGAT	CGCTGAGTAGGTG	4479
Qy	4571	cctcactgat	taagcattg	gttaact	gtcagac	ccaagttt	actcataatact	4630
Db	4480	CCTCACGTAT	TAAAGCAT	TGGTAAC	TGTCAGAC	CAAGTTTACT	CATATATACTTTAGATTG	4539
Qy	4631	atttaaaact	tcattttt	aaatttaa	aaagatct	aggtga	agatcctttt	4690
Db	4540	ATTTAAAC	TTTCATTTT	TATTTAA	AAGGATCT	AGGTGA	AGATCCTTTT	4599
Qy	4691	tgacaaaaa	ccccctaa	cgtagttt	tcgtcc	actgag	ctcagaccccg	4750
Db	4600	TGACCAAAAT	CCCCCTA	ACGTGAG	TTTTTCT	CCACTG	AGCGGTAG	4659
Qy	4751	tcaaaagat	ctcttgag	atcctttt	ctctgc	gcgtta	ctgcgtctg	4810
Db	4660	TCAAAGGAT	CTTCTTGAG	ATCCTTTTT	TCTGCGG	GTAACTCT	GCTGCTGCAACAAAAA	4719
Qy	4811	aacacacgt	ctacag	cggtgg	ttgttgc	cgga	tcaagag	4870
Db	4720	AACCACGCT	TACCAGCGGT	GGTTGT	TTTGC	GGATCA	AGAGTACCA	4779
Qy	4871	aggttaact	ggcttcac	gacgac	gagata	cacaaat	actgctc	4930
Db	4780	AGGTAACT	TGGCTTCAC	GACAGCG	CAGATAC	CAAAATAC	TGTCCTTCTAGT	4839
Qy	4931	taggccacc	acttca	agaactct	gtlag	cacgcctaca	ctccgtc	4990
Db	4840	TAGGCCACC	ACTTCAAGAA	CTGTGTAG	CACGCCCTAC	ATACCTCG	CTGCTAATCCTGT	4899
Qy	4991	taccagtg	gctgctgc	agtgcc	gataag	tcgtgt	ctta	5050
Db	4900	TACCAGT	TGGCTGCTGC	ACGTGG	CGATAAG	TGCTGT	TACCGGGT	4959
Qy	5051	agttacc	cgataa	ggcgcg	cggtcg	ggctga	cggggttc	5110
Db	4960	AGTTACC	CGGATAG	CGGCAC	ACGGTTC	GCGCTGA	ACCGGGGGTTC	5019
Qy	5111	tggagca	gaacga	actcac	gaactga	atac	ctacag	5170
Db	5020	TGGAGCAAC	AGACCTAC	ACCGAAT	CGATAC	CTACAC	GGTGAGCTAT	5079
Qy	5171	cgcttcc	gaaggga	aaaggc	gacaggt	atcc	ggtaag	5230
Db	5080	CGCTTCC	CGAAGGAG	AAAGGG	CGACAGAT	TCCGGT	TAAAGCGG	5139
Qy	5231	agccacg	agggag	cttc	cagg	ggaac	cgctg	5290
Db	5140	AGCCAC	CAGGGAGCT	TCCAGGG	GAACCG	CTGTATCT	TTTATAG	5199
Qy	5291	gccacct	ctgact	tgag	ctgatt	ttgtg	atgctc	5350
Db	5200	GCCACCT	CTGACT	TGAGCG	TCGATTTT	TGTGAT	GCTCTC	5259
Qy	5351	aaaacgc	acga	acgc	ggcctttt	acgg	tctgg	5410
Db	5360	AAAACGC	ACCAAC	GGGGCCTTTT	TACGGT	CTCTG	GGCCTTTT	5319
Qy	5411	tgcttct	ctcggt	tatcc	ctg	atctgt	gata	5470
Db	5320	TGTTCTT	CTCGGTTAT	CCCCCT	GTATTC	CTGTGG	ATAACCG	5379
Qy	5471	ctgata	ccgctc	gcgcg	agcc	gaacgc	gacgcg	5530

[illegible]

|||||
Db 5021 GCATGTGTCAGAGGTTTTACCGTCTATCACGAAACGGCGGAGACGAAGGCCCTCGTGA 5080
Qy 3551 tacgctattttttagattaatgcatgataataatggtttcttagcgtcagtgga 3610
Db 5081 TACGGCTATTTTTATAGGTTAATGTCATGATAAATAATGGTTTCTTTAGACGTCAGGTGGCA 5140
Qy 3611 cttttcggggaaatgctgcgggaacccctatttggttattttcttaataataataaata 3670
Db 5141 CTTTTCGGGGAAATGTCGGGAAACCCCTATTGTGTTTATTTTCTTAAATACATTCAATA 5200
Qy 3671 tqtatccgctcatgagacaataaacctctgataaattcttcaataatattgaaaaagaaaga 3730
Db 5201 TGTATCCGCTCATGAGACAAATAACCTTGATAAATGCTTCAATAATATGAAAAAGGAAGA 5260
Qy 3731 gtatgagatttbaacatttccgtgtcgcccttattcccttttttgcggcaattttcccttc 3790
Db 5261 GTATGAGTATTCAACATTTCCGCTGCGCCCTTATTCCTCTTTTTCGGGCATTTTGCCTTC 5320
Qy 3791 ctggtttttgctcaccagaaacgctgggtgaagtaaaagatgctgaagatcagttgggtg 3850
Db 5321 CTGTTTTTGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 5380
Qy 3851 cagcagtggtttacatcgaaactggatctcaacagcggtaagatcccttgagagttttcgcc 3910
Db 5381 CACGAGTGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC 5440
Qy 3911 ccgaagaacggttttccaatgatgagcaacttttaaaagttctgctatgtggtcggtattat 3970
Db 5441 CCGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGTATGTGGCGCGGTATTAT 5500
Qy 3971 cccgtattgagccgggcaagagaactcgtgcgcgcgatacaactattctcaagaatgact 4030
Db 5501 CCGGTATTGACCGCGGGGAAGAGCAACTCGGTGCGCGCATACACTATTCTCGAATGACT 5560
Qy 4031 tggttgagtactcacacagtcacagaaaagcatcttcaggatggcatgacagtaagagaat 4090
Db 5561 TGGTTGAGTACTCACAGTCACAGAAAGCACTTTACGGATGGCATGACAGTAAGAGAAT 5620
Qy 4091 tatgcaagtctgccataaccaatgadtgataacacactgcggcccaacttaactctgacaaaga 4150
Db 5621 TATGAGTGTCTGCCATAACCATGAGTGTATACACTTCGCGCCCAACTTACTTCTGACAACGA 5680
Qy 4151 tcggaggaccgaagagataaccgcttttttgcaacaatggggatcatgtaactcgcc 4210
Db 5681 TCGGAGGACCGAAGAGCTTAACCGCTTTTTCGCAACAATGGGGGATCATGTAACTCGCC 5740
Qy 4211 ttgacgtttgggaacggagctgaatgaagccataccaaacgacgagcgtgacacaaga 4270
Db 5741 TTGATCGTTGGCAACGGGAGCTGAATGAAGCCATACCAAAACGACGAGCTGACACCAGA 5800
Qy 4271 tgcctgtagcaatggcaacaacgcttgcgaacaactttaactggcgaaacttaactcttag 4330
Db 5801 TGCCCTGTAGCAATGGCAACAACGCTTCGCAAACTATTAACCTGGCGAACTACTTACTTAG 5860
Qy 4331 cttcccggaacaataatagatagctgagcggaacggaataaagtctcaggacacactctgc 4390
Db 5861 CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTTCAGGACCACTTCTGC 5920
Qy 4391 gtcggcccttcggctggctggtttattgctgataaaactctgagcccggtgagcgtgggt 4450
Db 5921 GCTCGGCCCTTCCGGCTGGCTGGTTATTTGCTGATTAATCTCGACCGGCTGAGCGTGGGT 5980
Qy 4451 ctcgcggtatcatctgacgactggggccagatggtaagccctcccgtatcgtagttatct 4510
Db 5981 CTCCGGGTATCATTCAGACACTGGGGCCAGATGGTAAGCCCTCCCGCTATCGTATATCT 6040
Qy 4511 acaacagcggggagtcaggcaactatggtgaacgaaatagacagatcgtctgagatggtg 4570
Db 6041 ACACACCGGGAGTCAGGCAACTATGTGATGAACGAATAAGACAGATCGCTGAGTAGGTG 6100
Qy 4571 cctcactgattaaagcattggtaactgtcagaccaagtttactcatatataactttagattg 4630
|||||

Db 6101 CCTCACTGATTAAGCATTTGGTAACTGTCTAGACCAAGTTTACTCATATATACATTTAGATTG 6160
Qy 4631 atttaaaacttcatttttaatttaaaagatctaggtgaagatcccttttttgataatctca 4690
Db 6161 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCA 6220
Qy 4691 tgacaaaatcccttaagctgaagttttggttcactagcgtcagaccccgtagaaaaaga 4750
Db 6221 TGACCAAAATCCCTTAAGCTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA 6280
Qy 4751 tcaaaaggaacttctcttgagatccctttttctcggtgaatctgctgcttgcaaaaaaaa 4810
Db 6281 TCAAAGGATCTCTTTGAGATCCCTTTTTCGCGGTAATCTGCTGCTTGCATAACAAAA 6340
Qy 4811 aaccaccgtaccagcgggtgtgtttgttgcgggatacaagagatcaacaactctttttccga 4870
Db 6341 AACCAACCGCTACACGCGGTGTTGTTTTCGCGGATCAAGAGCTACCAACTCTTTTCCGA 6400
Qy 4871 agttaactggcttcagcagagcgcagatacaaaaactgtccttcttagttagcgtagt 4930
Db 6401 AGGTAACCTGGCTTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTAGTGTAGCGGTAGT 6460
Qy 4931 taggcacaccacttcaagaactctgtagaccgcctacataacctcgtctgctlaactctgt 4990
Db 6461 TAGGCCACCACCTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCTGT 6520
Qy 4991 taccagtggtcgtccagtgccagataaagtcgtcttaccgggttggaactcaaacagat 6505
Db 6521 TACCAGTGGCTCTGCCAGTGGCGATAGTCGTGCTTACCGGGTGGACTCAAGACGAT 6580
Qy 5051 agttaccgggataaggcgcagcgggtcgggtgaacgggggttcgtgcacacagccagct 6640
Db 6581 AGTTACCGGATAAGCGCGACGCGTGGGTGAACGGGGGTTTCGTGCACACAGCCAGCT 6640
Qy 5111 tggagcgaacgacctcacaccgaactgagataacctacacgctgagctatgagaaagcgca 6700
Db 6641 TGGAGGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGGCCCA 6700
Qy 5171 cgtctcccgaaagggaaggcggacaggtatccggtaagcggcagggttcggaacaggag 6730
Db 6701 CGCTTCCCGAAGGGAAAGGGCGACAGGTATCCGTAAGCGCAGGGTTCGGAAACAGGAG 6760
Qy 5231 agcgacagaggagcttccagggggaacgcctggtatcttattatagctcgtcggttttc 6790
Db 6761 AGCGCACGAGGAGCTTCCAGGGGGAACGCTGTGTTCTTATAGTCTCTGCTGGGTTC 6820
Qy 5291 gcaacctctgacttgagcgtcatttttgatgctcgtcagggggggcggagcctatgga 6880
Db 6821 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTGCTGTCAGGGGGCGGAGCCATGGA 6880
Qy 5351 aaaaacgcagcaacgcggccttttttacgggttccctggccttttgcgtggccttttgcacaa 6940
Db 6881 AAAAGCCCAACGACGCGCTTTTACGGTTCTGTCGGCTTTTGTGCGCTTTTGTGCTCA 6940
Qy 5411 tgttttctcgtcggttatccctgattctgtgataaacgattaccgcttttgagtgag 6970
Db 6941 TGTCTTCTCGGTATATCCCTGATTCTGTGTGATTAACCGTATTACCCTCTTTGAGTGAG 7000
Qy 5471 ctgataccgctcgccgcagccgaacgacgagcagcagcagtcagtcagcaggaagcgg 7060
Db 7001 CTGATACCGCTCGCCGACGCCCAACGACCGACGCGAGCGATCAGTCAGCGAAGAACGG 7060
Qy 5531 aagagcgcccaatacagcaaacccgctctcccgcgcgtttggccgattcaatgagct 7120
Db 7061 AAGAGCGCCCAATACGAAACCGCTCTCCCGCGCGTTTGGCGGATTTCATTAATGACGCT 7120
Qy 5591 ggcagcaaggtttcccgaactggaagcggcagtgagcgaacgcaataatgtgagtt 7180
Db 7121 GGCACGACAGGTTTCCGACTGGAAGCGGGGAGTGAGCGCAACGCAATTAATGTGAGTT 7180
Qy 5651 agctcactcattagcaccacccagctttacactttatgcttccgctcgtatgttg 7240
Db 7181 AGCTCACTCATTTAGGACCCCGAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTG

QY 5711 gaattgtgagcgatacaaatccacacaggaacacgactatgaccatgattacgcca 5767
|||||
Db 7241 GAATTGTGAGCGGATAACAATTTCCACAGGAAACAGCTATGACCATGATTACGCCA 7297

RESULT 7
LOCUS AX114872/c 8062 bp DNA 11linear PAT 11-MAY-2001
DEFINITION Sequence 32 from Patent WO0129208.
ACCESSION AX114872
VERSION AX114872.1 GI:14031814
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1 (bases 1 to 8062)
AUTHORS Kuehn,R., von Melchner,H. and Altschmied,J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 32 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
Location/Qualifiers
1..8062
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pRK76"
BASE COUNT 1960 a 2036 c 2121 g 1945 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 8062;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaactggcgcgtgttttacaacgctgactgggaacacccctggcgttaacccaac 3190
Db 8045 CGAATTCACTGGCCGCGTGTGTTTACAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAAC 7986

QY 3191 ttaatgccttgacgacatccccccttttcgcaactgctggcgttaacgagagcccgca 3250
Db 7985 TTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATACGGAAGGCCCGCA 7926

QY 3251 cegatgccttcccaacagttgacgagcctgaatggcgaatggcgcctgacgagcgtatt 3310
Db 7925 CCGATCGCCTTCCCAACAGTTGCGCAGGCTGAAATGGCAATGGCGCCTGATCGGTATT 7866

QY 3311 ttctcttacgcatctgctggtatttcacaccgcatatggcactctcagtaacaatct 3370
Db 7865 TTCTCCTTACGCATCTGCGGGTATTTCACCCGCATATGGTGCACTCTCAGTACAATCT 7806

QY 3371 gctctgatgcgcgcatagtttaagccagcccgacaccccgcaacacccgctgacgcccct 3430
Db 7805 GCTCTGATGCGCATAGTTAAGCCAGCCGACACCCCGCAACACCCGCTGACGCGCCCT 7746

QY 3431 gacggctgtctgtcccggaatccgcttaacagacaagctgtgacggtctccggagct 3490
Db 7745 GACGGCGTGTCTGTCGCCGGCATCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 7686

QY 3491 gcatgttcagaggttttcacgctcatcaccgaacacgacgacgacgaagggcctcgtga 3550
Db 7685 GCATGTCTAGAGGTTTTCACCGCTCATCACCAGAACCGCGACGAAAGGGCCCTCGTGA 7626

QY 3551 taagcctattttatagggttaatgctatgataataatggtttcttttagacgctcaggtggca 3610
Db 7625 TAGCGCTATTTTATAGGTTAATGTCATGATATATATGTTTCTTAGACGTCAGGTGGCA 7566

QY 3611 cttttcggggaaatgtcgcgggaacccctattgtttattttttcttaaaacacattcaata 3670
Db 7565 CTTTTCGGGGAATGTGCGCGGAACCCCTATTGTTTATTATTTTCTTAATACATTCAATA 7506

QY 3671 tgtatccgctcatgacacaataaccctgataaatgctccaataattgaaaaagaga 3730
Db 7505 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA 7446

QY 3731 gtatgagtattcaacatttcgctgtcgcccttattcccttttttttggcgattttgcttc 3790
Db 7445 GTATGAGTATTCAACATTTCCGFGTCGCCCTTATTCCTTTTGGGCATTTTGGCTTC 7386

QY 3791 ctgttttttgcctcaccagaaacgctggtgaaagctaaagatcgtgaagatcgtgtgggtg 3850
Db 7385 CTGTTTTTGTCTACCCAGAAACGCTGCTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG 7326

QY 3851 cacgagtgggtttacatcgaactggatctcaacagcggtaagatccttgagagtttcgcc 3910
Db 7325 CACGAGTGGGTTCATCTGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTCGCC 7266

QY 3911 ccgaagaacgcttttccaatgatgagcaccttttaagttctctgctatgtggcggtattat 3970
Db 7265 CGGAAGAAGCTTTTCCAATGATGAGCACITTTAAAGTTCTGCTATGTGGCGGCTATTAT 7206

QY 3971 cccgtattacgcgcgggcaagagcaactcggctgcgcgcaatacactattctcagaatgact 4030
Db 7205 CCGTATTGACGCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACT 7146

QY 4031 tgggtgagtactcaccagtcacagaaagacatcttaacgatggcatgacagtaagaat 4090
Db 7145 TGGTTGAGTACTCACCAGTCCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAAT 7086

QY 4091 tatgcagtctgcataaaccatgagtataaacactcggcgcaacttactctctgacaacga 4150
Db 7085 TATGCAGTGTGTCATACCATGAGTGATAACACTCGGGCCAACTTACTTCTGACAACGA 7026

QY 4151 tcggagacgcgaaggactaacccgctttttgcaacaactgggggatactgtaactcgcc 4210
Db 7025 TCGGAGGACCGAGGAGCTAAACCGCTTTTGTGCACAACATGGGGGATCATGTAACTCGCC 6966

QY 4211 ttgatcgttgggaacgcgagctgaatgaagccatatacaaacgacgagcgtgacacacga 4270
Db 6965 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCCATACCAACGACGAGCGTGACACCACGA 6906

QY 4271 tgctgtagcaatggcaacaacgcttgcgaactatttaactggcgaactacttactcttag 4330
Db 6905 TGCTGTAGCATGTCACAACGTTGCGCAACTATTAACTGGCGAATCTACTTACTCTAG 6846

QY 4331 cttcccggaacaaatataagactggatggagcggaataaagtctgagacacactctgc 4390
Db 6845 CTTCCCGGCAACAATTAATAGACTGGATGAGCGCGATAAAGTTGCAGACCACTTCTGC 6786

QY 4391 gctcgcccttcgcgctgctggtttatgtgataaactcgtgagcgcgtgagcgtgggt 4450
Db 6785 GCTCGCCCTTCCGGCTGGCTGTTTTATTTGCTGATTAATCTGGAGCCGGTGACCGTGGGT 6726

QY 4451 ctgcgggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatct 4510
Db 6725 CTCGGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6666

QY 4511 aacgcggtggagtcaggcaactatggatgaacgaataagacagatcgcgtgagatagggtg 4570
Db 6665 ACACGACGGGGAGTCAGGCAACTATGATGAACGAATAACAGAGATCGCTGAGATAGGTG 6606

QY 4571 cctcactgattgaacattgttaactcagacccaagtttactcatatatactttagattg 4630
Db 6605 CCTCACTGATTAGCAATTTGGTAACCTGTACAGCCCAAGTTTACTCATATATACTTTAGATTG 6546

QY 4631 atttaaaactcttttttaatttaaagatctagtggaagatccctttttgtaaatctca 4690
Db 6545 ATTATAAACTTCATTTTAAATTTAAAGGATCTAGTGGAAGATCCCTTTTGTGATAATCTCA 6486

QY 4691 tgacaaaatcccttaacgctgagtttcttcacbtgcgctcagaccocgtagaagaaga 4750
Db 6485 TGACCAAAATCCCTTAACGTGAGTTTTGTTTCCACTTCAGCTGAGCGTCAGACCCCGTAGAAAAGA 6426

QY 4751 tcaagagctctctgagatcctttttctgcgcgtaactctgctctgcaacaaaaa 4810
Db 6425 TCNAAGATCTCTTGAGATCCCTTTTTTCTGCGCGCTAATCTCTCTCTGCAAAACAAAAA 6366

Qy	4811	aaccaccgtaccacggtgtgtttgtccgagatcaagagctaccaactctttttccga	4870
Db	6365	AACCACCCTACCAGCGTGTGTGTTCGGGATCAAGAGCTACCAACTCTTTTCCGA	6306
Qy	4871	aggtaacgtgcttcagcagcgcagataccaaaactactgtctctttagttagccgtagt	4930
Db	6305	AGGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCGGTAGT	6246
Qy	4931	taggcaccacttcaagaactctgtagcacgcgcctacatacctcgtctgtatacctgt	4990
Db	6245	TAGGCCACACTTCAAGAACTCTGTAGCAGCGCTACATACCTCTGCTGTGAATCTCTGT	6186
Qy	4991	taccagtgtcgtcgcagtgccagtaagtcgtgtcttaccgggttgacacaaagcat	5050
Db	6195	TACCAGTGGCTGCTGCCAGTGGCATAAGTGTGTCTTACCGGGTGGACTCAAGACCAT	6126
Qy	5051	agttaccggataaagcgcagcggctggcgtgaacggggggttcgtgcacacagccagct	5110
Db	6125	AGTTACCGGATAAGCGCAGCGGTGGGCTGAACGGGGGTTCGTGCACACAGCCAGCT	6066
Qy	5111	tggagcgaaacactacacccgaactgagatacctcacagcgtgagctatgaaaaagcgca	5170
Db	6065	TGGAGCGAACACCTTACACCGAACCTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCA	6006
Qy	5171	cgtctcccgaaaggagaaagcgcagcaggtatccggttaagcgcaggtcgaaacagag	5230
Db	6005	CGTCTCCGAAAGGGAAGAGCGGACAGGTATCCGGTAAAGCGGAGGTGCGAACAGGAG	5946
Qy	5231	agcgcagaggagcttccaggggaaacgcctggtatctttatagtcctgtcgggttcc	5290
Db	5945	AGCGACAGGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGTCTCTGTCGGGTTC	5886
Qy	5291	gccacctctgactgagcgtcgatttttgatgctcgtcagggggggcggagcctatgga	5350
Db	5885	GCCACTCTGACTTGAGCGTCGATTTTGTGTATGCTGCTCAGGGGGCGGAGCCTATGGA	5826
Qy	5351	aaaaagccagcaagcggcctttttacaggttccctggttcctgttcttggtgctttgtctaca	5410
Db	5825	AAAAGCCAGCAACCGGCGCTTTTACGGTTCCTGGCCCTTTTGTGGGCTTTTGTCTACA	5766
Qy	5411	tgtttttctcgttatccccctgattctgttggaataacgtattaccgcctttgagttag	5470
Db	5765	TGTTCTTTCCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCTTTGAGTGAG	5706
Qy	5471	ctgataccgcgtcgcgcgcgcgaacgcagcagcgcagcagtcagtcagtcagcaggaagcgg	5530
Db	5705	CTGATACCGCTCGCGCGACCGCGAACGACGACGCGCAGCGAGTCAGTGAGCGAGGAAGCGG	5646
Qy	5531	aagagcgcccaatagcaaaacgcctctcccgcgcgcttgccgattcattaatcagct	5590
Db	5645	AAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCAATATGACGT	5586
Qy	5591	ggcacgacaggtttcccgactggaaagcggcagtgagcgaacgcaatattgtagtt	5650
Db	5585	GGCACGACAGGTTTCCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTT	5526
Qy	5651	agctcaactaattaggcacccagcgttttacactttatgcttccggctcgtatgtgtgtg	5710
Db	5525	AGCTCACTCATTAGCGACCCAGGCTTTACACTTTATGCTTCGGCTCGTATGTTGTGTG	5466
Qy	5711	gaattgtgagcggataaacaatttccacaggaagcaagcatatgacctattacgcga	5767
Db	5465	GAATTGTGAGCGGATAAACAATTTACACAGGNAACAGCTATGACCATGATTACGCCA	5409
RESULT 8			
AX114871/c			
LOCUS	AX114871	8153 bp	DNA linear PAT 11-MAY-2001
DEFINITION	Sequence 31 from Patent WO0129208.		
ACCESSION	AX114871		
VERSION	AX114871.1 GI:14031813		
KEYWORDS	synthetic construct.		
SOURCE			

ORGANISM	synthetic construct		
REFERENCE	1 (bases 1 to 8153)		
AUTHORS	Kuehn,R., von Melchner,H. and Altschmied,J.		
TITLE	Conditional gene trapping construct for the disruption of genes		
JOURNAL	Patent: WO 0129208-A 31 26-APR-2001.		
FEATURES	ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)		
source	Location/Qualifiers		
	1..8153		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="vector PRK7"		
BASE COUNT	1976 a	2133 c	2067 g 1977 t
ORIGIN			
Query Match	45.7%	Score 2637;	DB 6; Length 8153;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2637;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	3131	cgaattcactggccgtcgtttttacaacgtcgtgactgggaaaaacctgcttaccacaac	3190
Db	8136	CGAATTCACTGGCCGTCGTTTACAACTGCTGACTGGGAAACCTGGCGTTACCCAAC	8077
Qy	3191	ttaatcgctctgcagcacacatccccctttccagctggcgtaataagcgaagagccgcga	3250
Db	8076	TTAATCGGCTTGCACACATCCCCCTTTCGCCAGCTGGCGTAAATAGCGAAGAGGCCGCA	8017
Qy	3251	ccgactcgccctcccaacagttgcgcagctgaatggcgaatggcgctgagcggtatt	3310
Db	8016	CGATCGCCCTTCCCAACAGTTGCGCAGCGCTGAATGGGAAATGGCGCTTGATGCGGTATT	7957
Qy	3311	ttctccttacgcactgtgcggtatttcacacccgcataatggcgactctcagtcacaact	3370
Db	7956	TTCTCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCATCTCTCACTACAACT	7897
Qy	3371	gctctgacgcgcatagtttaagcggcccgccgcacacccgcgcgcgcgcgcgcgc	3430
Db	7896	GCTCTGATGGCGCATAGTTAAAGCCAGCGCCGACACCCGCCAACACCCGCTGACGGCCCT	7837
Qy	3431	gacgggctgtctgtcccgccgcacccgtctacagacaagctgtgacctctccggagct	3490
Db	7836	GACGGGCTTGTCTGCTCCCGGCATCCGTTACAGACAAGCTGTGACCGTCTCCGGAGCT	7777
Qy	3491	gcattgtcacagaggttttcacccgtcatcacgaaaacgcgcgcgcgcgcgcgcgc	3550
Db	7776	GCATGTGCAGAGGTTTTCACCGTCACTCACCGAAACGGCGAGACGAAAGGCCCTCGTGA	7717
Qy	3551	tacgcctattttataggttaatgctcatgataataatggttctcttagacgtcaggtggca	3610
Db	7716	TACGGCTATTTTATAGGTTAATGTCATGATAAATGGTTTCTTAGACGTCAAGTGCCA	7657
Qy	3611	cttttcgggggaaatgctgcgggaacccctattgtttattttctcaaatacattcaata	3670
Db	7656	CTTTTCGGGGAATGTGCGCGAACCCCTATTTTGTATTATTTTCTAATACTACAAATA	7597
Qy	3671	tgatccgcctcatgagacaataaacccctgataaattcttcaataattgaaaaagaaga	3730
Db	7596	TGATCCGCTCATGAGCAATAAACCTGATAAATGCTTCAATAATATGAAAAAGAAAGA	7537
Qy	3731	gtatagatttcaacatttccgctgcgccttatcccccttttttttcggcatttgccttc	3790
Db	7536	GTATAGATTTCACATTTCCTGCGCCCTATTCCCTTTTTCGGCATTTTTCCTTC	7477
Qy	3791	ctggtttttgctcaccagaaaacgctggtagaagtaaaagatgctgaagatcagttgggtg	3850
Db	7476	CTGTTTTTGTCTACCCAGCAAAACCCCTGGTGAAAGTAAAGATGCTCAAGATCACTTGGG	7417
Qy	3851	cacgagtggtttacatcgaactggatctcaacagcggtaagatcctttagagattttcgcc	3910
Db	7416	CACGAGTGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGATTTTCGCC	7357
Qy	3911	ccgaagaaacggtttttccaatgatgagcaacttttaaaagttctgtctgtggtggtggtatt	3970

JOURNAL Submitted (05-JUN-1995) Paul A. Kitts, CLONTECH Laboratories, Inc., 4030 Fabian Way, Palo Alto, CA 94303, USA
COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 4030 Fabian Way, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3. This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.

FEATURES Location/Qualifiers
source 1..4518
/organism="Cloning vector p35S-GFP"
/plasmid="p35S-GFP"
/db_xref="taxon:41037"
896..1612
/gene="gfp"
896..1612
/gene="gfp"
/codon_start=1
/transl_table=1
/product="green fluorescent protein"
/protein_id="AAA59675.1"
/db_xref="GI:894140"
/translation="MSKGEELFTGVPIVLYFGNPILVDESDVNGHKFVSQGESEGDATYGLKLTKFICTGKLPVPTLVFTTYSYVQVCFSRYPDHMKQHDFFKSPAMPEYVQDRTFFKDKVNYKRAEVKFGDGLVNRLELKGIDFKEDGNILGHLENYNHNHYIMADKQKNGIKNFKIRHNIEDSVOLADHYQNTPIGDGPLLPDNLHSTOSALSKEPDNEKRDEHVLLEFVTAAGTHGMDLYK"
2485..3345
/gene="bla"
2485..3345
/gene="bla"
/function="confers resistance to ampicillin"
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="AAA69676.1"
/db_xref="GI:894141"
/translation="MSIQHPRVALIPFFAAFLPFAHPETLVKVKDAEDOLGARVGYIELDLSGKILESFRPERPMSSTFKVLCGLVLRIDAGQEOVGRIHYSQNDLVEISPYERKHLTDGMTVRELCASAIITMSNTLANLLLTIGGPKELTAPLHNMGDHYTRLDRWPELNEAIPNDRDITMPVAMATTIRKLTLTGELTLASRQQLIDMMEADKVGPLLRSLAPAGWFTADSKSGAGERSGRIIAALGPDGPKSRIVLYITGSOATMDERNRQIAEIGASLIKHW"
BASE COUNT 1287 a 1055 c 1042 g 1134 t

Query Match 45.7% Score 2635; DB 12; Length 4518;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3132 gaattcaatgcccgtcggttttacaacgtcgctgactggagaaacccctggcttacccaact 3191
|||||
Db 1884 GAATTCACGTGGCGCGTCTTTACACAGTCGTGACCTGGGAAAAACCCCTGGCTTACCCAACT 1943
Qy 3192 taatgcctttgcagacacatcccccttcctgcagctggtcggtgtaagcggaatggcgccctgactgacggtattt 3251
|||||
Db 1944 TAATCGCCTTGAGACATCCGCCCTTTTCGCCAGCTGGCGGTAAATCGAAGAGGCCCGCAC 2003
Qy 3252 gtagcgccttcccacagttgcgcagcctgaatggcggaatggcgccctgactgacggtattt 3311
|||||
Db 2004 CGATCGCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCCTGATGCGGTATT 2063
Qy 3312 tctcttaacgactctgtcggttatctaacacgcgcatatggtgacactcactgacacactg 3371
|||||
Db 2064 TCTCCTTACGCATCTGTGCGGTATTTACACCGGCATATGGTGCATCTCTCAGTACAATCTG 2123
Qy 3372 cctctgacgcgcatagttaagccagccccgcacacccgcgcaacccgctgacgcgccttg 3431

Db 2124 CTCATGCCGATAGTTAAGCCAGCCCGACACCCGCCAACACCCGCTGACGCCCTCG 2183
Qy 3432 acgggcttctgtctcccgccgcatccgtctacagacaagctgtgacccgtctccggagctcg 3491
|||||
Db 2184 ACGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGAGCTG 2243
Qy 3492 catgtctcagaggttttccacogtcatcccgaaacgcgcgagacgaaaggcctcgtag 3551
Db 2244 CATGTGTGAGAGGTTTTCACCGTCAACCGAAACCGCGAGACGAAAGGCGCTCGTGTAT 2303
Qy 3552 acgctatttttataggtaattgtcatataataatggttttcttagacgtcaggtggcac 3611
|||||
Db 2304 ACGCCTATTTTATAGGTTAATGTTCATGATAATAATAGTTTCTTAGACGTACAGTGGCAC 2363
Qy 3612 ttttcgggaaatgtgcgcggaacccctatttggttatttttcttaataacattcaaatat 3671
Db 2364 TTTTCGGGGAATGTGCGGGNAACCCCTATTGTTTATTTTCTAATACATTCAAATAT 2423
Qy 3672 gtatccgctcatgagacaataaccctgataaatgtctcaataattgaaaaaggaagag 3731
Db 2424 GTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATTTGAAAAAGGAAG 2483
Qy 3732 tatgagattcaacatttccgtgtcgcccttattcccttttttggggcaatttgccttcc 3791
Db 2484 TATGAGTATTCAACATTTCCGCTGTCGCCCTTATTCCCTTTTTCGGGCATTTTGCCTTC 2543
Qy 3792 tgttttctcacccagacaacgcgtgtgaaagtaaaagatgctgaagatcaggttgggtgc 3851
Db 2544 TGTTTTTCTCACCCAGAAACGCTGGTGAAGTAAGAATGCTGAAGATCAGTTGGGTGC 2603
Qy 3852 acgagtggtttacatcgaaactggatctcaacagcggtgaagatccttgagagtttccgcc 3911
Db 2604 ACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCTTGAGAGTTTTCGCC 2663
Qy 3912 cgaacacgctttccaatgatgagcaattttaaagttctgctatgtggcggtattatc 3971
Db 2664 CGAAGACGTTTTCCAAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCGTATTATC 2723
Qy 3972 ccgtattgacgcgggcaagagcaactcggctgcgcgcacacactattctcagaatgactt 4031
Db 2724 CGGTATTGACGCCGGCAAGACCACTCGGTGCGCGCATACACTATTCTCAGAAAGACTT 2783
Qy 4032 ggttgacttaccagtcacagaaaagatcttaccgagtcgacagtaagaagaatt 4091
Db 2784 GGTTCAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGCGATGACAGTAAGAATA 2843
Qy 4092 atgcagctgcccataaacatgagtgataaacactcgggccaacttacttctgacaacgat 4151
Db 2844 ATGCAGTGTGCCCATAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTGACACGAT 2903
Qy 4152 cggaggaacgagagctaacccgctttttgcacaacatgggggagatcgttaactcgcct 4211
Db 2904 CGGAGGACCGAAGAGACTAACCGCTTTTTCGCAACATGCGGGATCATGTAACTCGCT 2963
Qy 4212 tgatcgttgggaacccgagctgaaatgaagccatcacaaacgcagcagcgtgacacacgat 4271
Db 2964 TGATCGTTGGGAACCGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACCACT 3023
Qy 4272 gctctagcaatggcaacacgcttgcgcaacatataactggcgcaacttacttctagc 4331
Db 3024 GCCTTAGCAATGGCAACAACCGTTCGCCAACTATTTAACTGGCGCACTACTTACTCTAGC 3083
Qy 4332 ttcccgggcaacaataatagatgagtgagcggaataaagttgcaggaacccctctgcg 4391
Db 3084 TTCCCGCGCAACAATTAATAGACTGGATGGAGCGGGGATAAAGTTGCAGGACCACTTCTCGG 3143
Qy 4392 ctccgaccttccgctggctggttttatgtgataaaatctggagccggtgagcgtgggtc 4451
Db 3144 CTCGCGCCTTCGCGCTGGCTGCTTTATTGCTGATAAATCTGAGCGCGGTGAGCGTGGGTC 3203
Qy 4452 tcgcggtatcatgtcagacactggggccagatggtaagccctccccgtatcgtagtattcta 4511
|||||

Db 3204 TCAGGCTATCATTCACCACTGGGGCCAGATGGTAACCCCTCCGGTATCGTAGTTATCTA 3263
Qy 4512 cagacggggagtcagcaactatgagatgaacgaaatagacagatcgctgagatagtc 4571
Db 3264 CAGCAGGGGAGTCAGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGTGC 3323
Qy 4572 ctacgtgattaagcattggttaactgtcagaccaaagtttaactatataactttagattga 4631
Db 3324 CTCACATGATTAAAGCATGTGTAACGTGCAGACCAAGTTTACTCATATACTATTAGATTGA 3383
Qy 4632 ttcaaaacttcatttttaatttaaaaggatctaggtagaagatccttttgaataatcctat 4691
Db 3384 TTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATATCTCAT 3443
Qy 4692 gacaaaatcccttaactgagtttctgtccactgagcgtcagaccccgtagaaaaagat 4751
Db 3444 GACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGCTAGAAAAGAT 3503
Qy 4752 caaaggatctcttgagatccttttttctgtgcggttaactgtcgttgcgaacaaaaa 4811
Db 3504 CAAAGGATCTTCTTGAGATCCTTTTTCGCGGTAACTGCTGCTTGCAAAACAAAAA 3563
Qy 4812 accacgctaccagcgggtgtttgttcgcggtcaagagctaccactctttttccgaa 4871
Db 3564 ACCACGCTACCAAGCGTGTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCGAA 3623
Qy 4872 ggtactgggttcagcagagcgcagataccataactgtccttcttagtgcgctagtt 4931
Db 3624 GGTAACTGGCTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTCTAGTGTAGCGTAGTT 3683
Qy 4932 agccaccacttcaagaactctgtagcagcgcctacataactcgtcgtctgtaactctgtt 4991
Db 3684 AGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGTGTAATCTCTGT 3743
Qy 4992 accagtgtctgtcgcagtgagcgaatagtcgtcttaccgggttgagactcaagacata 5051
Db 3744 ACCAGTGGCTGTGCCAGTGGCGATAACTGCTGTCTTACCGGTTGGACHTCAAGACGATA 3803
Qy 5052 gttaccggataaaggcagcgggtcgggctgaacgggggttcgtgcacacagccagctt 5111
Db 3804 GTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCGTGCACACAGCCCCAGCTT 3863
Qy 5112 gtagcgaacgacctacacgaactgagatgatacctacagcgttgagactatgagaagcgccac 5171
Db 3864 GGAGCGAAGCACTACACCACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAC 3923
Qy 5172 gcttcccgaaaggagaaaggcgacaggtatccggttaagcggcagggctcggaacagagaga 5231
Db 3924 GCTTCCGAAAGGAGAAAGCGCACAGGTATCCGGTAAGCGCGCAGGGTCGGAACAGGAGA 3983
Qy 5232 ggcacagggagcttccagggggaacgcctggtatctttagtctctcggggttcg 5291
Db 3984 GCGCAGGAGGAGCTTCAGGGGGAACAGCGCTGTATCTTTATAGTCTCTCGGGTTTCG 4043
Qy 5292 ccacctgactgagcgtcgattttgtgatactcgtcaaggggggcgagcctatggaa 5351
Db 4044 CCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTAGGGGGCGGAGCCCTATGGAA 4103
Qy 5352 aaacgcagcaacgcggccttttaaggttcctggccttttgccttttgccttttgcctacat 5411
Db 4104 AAAGCCAGCAACCGCGCTTTTACGGTTCCTGGCCTTTTGTGCTGGCTTTTCTCATAT 4163
Qy 5412 gttcttctcgtgtatccctctgattctg tggataaccgctattaccgcctttgagtgagc 5471
Db 4164 GTTCTTTCTTCGCTTATCCCTCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGC 4223
Qy 5472 tgatacgcctcgcgcagcgaacgacccagcagcagcagcagcagcagcagcagcagcagcga 5531
Db 4224 TGATACCGCTCGCGCAGCGAAGACCGAGCGCAGCGAGTCTAGTGAGCGAGGAACGGA 4283
Qy 5532 agagcggccaaatcgaacacgcctctccccgcgcgttggccgattcattaatgacgctg 5591
Db 4284 AGAGCGCCCAATACGAAACCGGCTCTCCCGCGGCTTGCCCGATTCATTAATGCAGCTG 4343

Qy 5592 gcacgacaggtttcccgactggaaagcggcgagtgagcgcaacgcaataatgtgagtta 5651
Db 4344 GCACGACAGGTTTCCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTA 4403
Qy 5652 gctcactcattagcaccgccagcgttttaacattttatcttccgctcgtatgttgtgtg 5711
Db 4404 GCTCAGCTCATTAGGCACCCCGAGGCTTTTACACTTTATGCTTTCCGGCTCGTATGTGTGG 4463
Qy 5712 aattgtgagcggataacaatttcacacaggaacacagctatgaccatgattacgcc 5766
Db 4464 AATTGTGAGCGGATACAAATTTTACACAGGAACAGCTATGACCATGATTTACGCC 4518
RESULT 10
U02456
LOCUS U02456 4519 bp DNA circular SYN 29-MAR-1996
DEFINITION Cloning vector pRAJ275, complete sequence.
ACCESSION U02456
VERSION U02456.1 GI:413822
KEYWORDS
SOURCE Cloning vector pRAJ275.
ORGANISM Cloning vector pRAJ275
REFERENCE 1 (bases 1 to 4519)
AUTHORS Kitts, P.A.
TITLE ClONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4519)
AUTHORS Jefferson, R.A.
TITLE Plant reporter genes: the GUS gene fusion system
JOURNAL (in) Setlow, J.K. (Ed.);
GENETIC ENGINEERING: PRINCIPLES AND METHODS;
10: 247-263 (1988)
REFERENCE 3 (bases 1 to 4519)
AUTHORS Kitts, P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. This vector is no longer
available from CLONTECH and CLONTECH will not update or revise this
sequence.
FEATURES
source Location/Qualifiers
1..4519
/organism="Cloning vector pRAJ275"
/db_xref="taxon:31825"
BASE COUNT 1130 a 1131 c 1190 g 1068 t
ORIGIN
Query Match 45.7%; Score 2635; DB 12; Length 4519;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3132 gaattcactggcgcgtgcttttaacaacgtcgtgactgggaacccctggcgttaccaccaact 3191
Db 1885 GAATTCACCTGGCGTCTGTTTACAACTGCTGACTGGGAAACCCCTGGCGCTTACCCAACT 1944
Qy 3192 taatgccttgagcacatccccctttgcgcagctggcgttaataaggaagagccgcac 3251
Db 1945 TAAATGCGCTTGCAGACATACCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCAC 2004
Qy 3252 cgatgccttcccaacagcttcgcagcctgaatggcgaatggcctgattgcggtattt 3311
Db 2005 CGATCGCCCTTCCCAACAGCTTGGCGAGCCCTGAATGGCAATGGCGCTGATGCGGTATT 2064
Qy 3312 tctccttaagcatctgtggtgattttcacaccgcataatggtgcactctcagtaacaactg 3371
Db 2065 TCCTCTACGCATCTGCGGTATTTACACACCGCATATGTCACCTCTCTAGTACAACTG 2124
Qy 3372 ctctgtagcgcagtagttaagcagccccgcacccgcgacccgctgacgcgcctg 3431

|||||
Db 2125 CTCGATGCGGATAGTTAAGCCAGCCCGACACCCGCCCAACACCCGCTGACGCCGCTCG 2184
QY acgggcttgcctcccgccgcatccgtttacagacaagctgtgaaccgtctccggagctg 3491
|||||
Db 2185 ACGGGCTGTCTGCTCCCGGCATCCGTTACAGACAAGCTGTGACCGTCTCCGGGAGCTG 2244
QY catgtgtcagagggtttccaccgctcataccgaaacacgcgcagacagaaaggccctcgtgat 3551
|||||
Db 2245 CATGTGTGTCAGAGGTTTTCACCGTCATCACCGAAACCGCGAGACGAAAGGGCTCGTGAT 2304
QY acgcctattttataggttaattgtcatgataataatggtttcttagacgtcaggtggac 3611
|||||
Db 2305 ACGGCTATTTTATAGGTTATGTCTATGNTAATATATGTTTCTTAGACGTGAGGTGGCAC 2364
QY ttttcggggaattgtgcgcgaacccctattgttttatttttctaaatacatctcaaatat 3671
|||||
Db 2365 TTTTCGGGGAATGTGCGGGAACCCCTATTGTTGTTATTTTCTAAATACATTCAAAATAT 2424
QY gtatccgctcatgagacaataaccctgataaattgcttcaataattgaaaggaagag 3731
|||||
Db 2425 GTATCCGCTCATGAGACAATAACCCGTGATTAATGTTTCAATATATTTGAAAAGGAAGAG 2484
QY tatgagattcaaacatttcctgtcgccttattcccttttttgcgcattttgcctcc 3791
|||||
Db 2485 TATGAGTATTCACATNTTCCGTGTGCGCCCTTATTCCTTTTTTGGCGCATTTTGGCTTCC 2544
QY tgttttctcaccagaaacgctggtgaagtaaaagatgctgaagatcagttgggtgc 3851
|||||
Db 2545 TGTTTTTGCTCACCCAGAACGCTGGTGAAGTAAGATGCTGAAGATCAGTTGGGTGC 2604
QY acgagtggtttacatcgaaactggatctcaaacgcggtgaagatcccttgagagtttcgccc 3911
|||||
Db 2605 ACGAGTGGGTTACATTCGAACCTGGATCTCAACAGCGGTGAAGATCCTTGAGAGTTTTCGCC 2664
QY cgaagacggtttccaatgatgagcaccttttaagttctctgctatgtgcccgtattatc 3971
|||||
Db 2665 CGAAGACGTTTCCATGATGAGCACNTTAAAGTTCTGCTATGTGGCGCGGTATTATC 2724
QY ccgtattacgcgggggaaggaactcgggtgcgcgcatacactattctcagaatgactt 4031
|||||
Db 2725 CCGTATTACGCGCGGAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTT 2784
QY ggttgagtactcaccagtcacagaaaagcatcttcaggtatggcatgacaagagaatt 4091
|||||
Db 2785 GGTTCAGTACTCACCACTACAGAAAGCATCTTACGATGGCATGACAGTAAGAGAAAT 2844
QY atgcagtgctgcataaacatgagtgaataacactgcgcgaacttacttctgacacgat 4151
|||||
Db 2845 ATGCACTGCTGCCATTAACCATGATGATTAACACTGCGGCCAACTTACTTCTGACAAACGAT 2904
QY cggaggaacggaaggctaaaccgcttttttgacaacatggggggtcatgtaactcgcct 4211
|||||
Db 2905 CGGAGGACCGAAGGAGCTAACCGCTTTTGTGCACAACATGSGGGATCATGTAACTCGCT 2964
QY tgaatcgttgggaacggagctgaatgaagccatacaaacacgagcagcgtgacacacgat 4271
|||||
Db 2965 TGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCAACGAT 3024
QY gctgtagcaatggcaaacgcttgcgaactatttaactggcgaacttacttactcagc 4331
|||||
Db 3025 GCCTGTAGCAATGGCAACACGTTGCGCNAACATTTAACTGGCGAANTACTTACTCTAGC 3084
QY ttcctgggaacaaatttaatagactggtgagcggtgataaagttaggagaccacttctgcg 4391
|||||
Db 3085 TTCCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGAGGACCACTTCTCGG 3144
QY ctccgcccctccgctgctggtttatttctgataaactgagccggtgagcgtgggtcc 4451
|||||
Db 3145 CTCGCCCTCTCCGGTGGCTGCTGTTTATGCTGATAAATCTGAGCGCGGTGAGCGTGGT 3204
QY tcgcggtatcattgagcactgggcccagatggtgaagccctcccgatctcgtattatc 4511
|||||

Db 3205 TCGCGGTATCATTTGCAGACTCGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTA 3264
QY caccgagggagtcaggcgaactatgataaagaaatagacagatcgctgagataggtgc 4571
|||||
Db 3265 CACGACGGGAGTCAAGCAACTATGGATCAACAAATACACAGATCGCTGAGATAGTGC 3324
QY ctcaactgatgaagcattggtaactgtcagaccaaagtttactcatatatactatttagattga 4631
|||||
Db 3325 CTCACGTGATTAAGCATTTGGTAACCTGTACAGCAAGTTTACTCATATATATCTTTAGATTGA 3384
QY tttaaaaactctatttttaatttaaaaggtactaggtgaagatcccttttggataaactcat 4691
|||||
Db 3385 TTTAAACTTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTTTGATANTCTCAT 3444
QY gacaaaaatcccttaacgtgagtttcttccactgagcgtcagaccccgtagaagaat 4751
|||||
Db 3445 GACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTCAGCGTCAGACCCCGTAGAAAGAT 3504
QY caaagatctctttagagatcccttttttctgagcgtaaactctgctgcttgcacaaacaaaa 4811
|||||
Db 3505 CAAAGGATCTTCTTGAGATCCTTTTTTCTGCGGTAATCTGCTGCTGCAACACAAAAA 3564
QY accacgcctaccagcgtggtttgttgcgcgatacaagagctaccacactcttcttcgaa 4871
|||||
Db 3565 ACCACGCTACCGCGGTGTTGTTGCGGATCAAGAGCTACCAACTCTTTTTTCCGAA 3624
QY ggtaactggcttcagcagagcagataccaaatactgtccttcttagttagcgtagtt 4931
|||||
Db 3625 GGTAACTGGCTTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTT 3684
QY aggcacacacttcaagaactctgtagcaccgctacatacctcgtctcgtctaatcctggt 4991
|||||
Db 3685 AGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATCTGTT 3744
QY accagtggctgctgcgcagtgcgataagctgctgtcttaccggtgtgactcaagacgata 5051
|||||
Db 3745 ACCAGTGGCTGCTGCCAGTGGCGATAACTGCTCTTACCGGTGGACTCAAGACGATA 3804
QY gttaccggataaggcgcagcgtgcgggtgaacgggggtctcgtgcacacagcccagctt 5111
|||||
Db 3805 GTTACCGATAAGGCGCAGCGTTCGGGCTGAACGGGGGTTGCTGCACACAGCCAGCTT 3864
QY ggaggaacgcactacacacgaactgagataactacagcgtgagctatgagaaagcccac 5171
|||||
Db 3865 GAGCGGAACGACCTACACGAACCTGAGATACCTACAGCTGAGCTATGAGAAAGCCAC 3924
QY gttcccggaagggaagcgcgcagcgtatccggtgaagcgcagcgttcggaaacagaga 5231
|||||
Db 3925 GCTTCCGAAAGGAGNAAGCGCGACAGGTATCCGGTAAAGCGCAGGGTCCGAAACAGAGA 3984
QY gcgcagagggagcttccaggggggaacgccttggtatcttttagtctcgtcgggttcg 5291
|||||
Db 3985 GCGCAGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTCTGCTGGGTTTCG 4044
QY ccactctgaactgagcgtcgatttttgatgctcgtcagggggcgagcctatggaa 5351
|||||
Db 4045 CCACCTCTGACTGAGCGTCGATTTTGTGATCTCTCAGGGGGCGGAGCCTATGGAA 4104
QY aaacgcagcaacgcgcctttttagcgttctcgtcgttcttgccttcttgccttctcacat 5411
|||||
Db 4105 AAAGCCACACAGCGGCGCTTTTACGGTTCTTGGCCTTTTGTGCTGCTTCTGCTCCTCAT 4164
QY gttcttctcgtgattatcccttgattctgtgataaacctgattaccgcctttagtgagc 5471
|||||
Db 4165 GTTCTTCTCGCTTATCCCTGATTTCTGTGGATAACCGTATTTACCGCTTTTTCAGTGAGC 4224
QY tgataccgctcgcgcagcgcgaacacagcagcgcagcgtcagtgagcagagaaagcga 5531
|||||
Db 4225 TGATACCGCTCCCGCAGCCGAACCGCAGCGAGTCAAGTCAAGCGAGGAAAGCGGA 4284
QY agagcgcacaaatcacaaacccctctcccgcttgcgcgttgcgattcatttaagcgcgt 5591
|||||
Db 4285 AGAGGCCCAATACGCAAAACCCGCTCTCCCGCGGCTTGGCCGATTTCAATTAATGCAGCTG 4344


```
QY 5592 gacgacaggtttcccgactgaaagcgggagtgagcgaacgaatgaattgagttga 5651
|||||
Db 4345 GCACGACAGGTTTCCGACTGGAAAGCGGCGAGTGAAGCAATTAATGTGAGTTA 4404
QY 5652 gctcactcattagcagccagctttacactttatgcttccgctcgatgtgtg 5711
|||||
Db 4405 GCTCACTCATTAGCACCCAGGCTTTACACTTTATGCTTTCCGGCTCGGTATGTTGTGG 4464
QY 5712 aattgagcggataaattcacacagaaacagatgacatgacattacgcc 5766
|||||
Db 4465 AATTGTGAGCGGATAACAATTTACACAGGAACACGATATGACCATGATTACGCC 4519

RESULT 11
XXU02437
LOCUS          XXU02437          4674 bp      DNA      circular SYN 29-JAN-1997
DEFINITION     Cloning vector pT3/T7-LUC, complete sequence.
ACCESSION      U02437
VERSION        U02437.1  GI:413803
KEYWORDS       .
ORGANISM       Cloning vector pT3/T7-LUC.
SOURCE         Cloning vector pT3/T7-LUC
               artificial sequence; vectors.
REFERENCE      1 (bases 1 to 4674)
AUTHORS        Kitts,P.A.
TITLE          ClONTECH Vectors On Disc version 1.3
REFERENCE      2 (bases 1 to 4674)
AUTHORS        Kitts,P.A.
TITLE          Direct Submission
COMMENT        Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
               1020 East Meadow Circle, Palo Alto, CA 94303, USA
               This vector can be obtained from CLONTECH Laboratories, Inc., 1020
               East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
               call (415) 424-8222 or (800) 662-2566, extension 1. International
               customers, please contact your local distributor. For technical
               information, call (415) 424-8222 or (800) 662-2566, extension 3.
               This sequence has been compiled from information in the sequence
               databases, published literature and other sources, together with
               partial sequences obtained by CLONTECH; this vector has not been
               completely sequenced. If you suspect there is an error in this
               sequence, please contact CLONTECH's Technical Service Department at
               (415) 424-8222 or (800) 662-2566, extension 3 or E-mail
               TeCh@CLONTECH.COM.
FEATURES
   source       Location/Qualifiers
               1..4674
               /organism="Cloning vector pT3/T7-LUC"
               /db_xref="taxon:31048"
BASE COUNT     1239 a 1092 c 1141 g 1202 t
ORIGIN
```

Query Match 45.7%; Score 2635; DB 12; Length 4674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3133 aattcactggccgctgttttacacgctgctgactgggaaacccctggcttaaccaactt 3192
|||||
Db 2008 AATTCACTGGCCGTCGTTTTACAACTCGTGACTGGGAAACCCCTGGCGTTACCCAACTT 2067
QY 3193 aatcgcttcagacacatccccctttccgagctggcgtaataagcgaagggccgcaac 3252
|||||
Db 2068 AATCGGCTTCGACGACATCCCTCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCACC 2127
QY 3253 gatcgccctcccaacagttgacagcctgaatggcgaatggcgctgatgcggtatttt 3312
|||||
Db 2128 GATCGGCCCTTCCCAACAGTTGCCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTTATTT 2187
QY 3313 ctccctacgcatctgtgcggtatttcacacgcgcataatggcgactctcagtaaatctgc 3372
|||||
Db 2188 CTCTTACGCATCTGTGCGGTATTTTACACACCGCATATGGTGCACTCTCACTACAAATCTGC 2247
```

```
QY 3373 tctgatgcgcgcatagtttaacgacgccccgcgacacaccccgccacccgctgacgcccctga 3432
|||||
Db 2248 TCTGATGCCGCATAGTTAAAGCCAGCCCGACCGCAACCCGCTGACGCGCCCTGA 2307
QY 3433 cggccttgctgctcccgccatccgcttacagaaagctgtagccgctcccgagagctgc 3492
|||||
Db 2308 CGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGC 2367
QY 3493 atgtctcagagggtttccaccgtcatccaccgaaacgcgcgagagaaagggccctcgtgata 3552
|||||
Db 2368 ATGTGTCAGAGGTTTTCACCGTTCATCACCGAAACGCGAGAGAAAGGCGCTCGTGATA 2427
QY 3553 cgctatttttataagtttaagttaagttaagttaagttaagttaagttaagttaagttaagt 3612
|||||
Db 2428 CGCCTATTTTATAGGTTAATATGATGATGATGATGATGATGATGATGATGATGATGATG 2487
QY 3613 ttccgggaaatgctgcgcgaaacccctattgtttattttttctaaataacattcaaaatg 3672
|||||
Db 2488 TTTCCGGGAAATGTGCGGGAACCCCTATTGTGTTTATTTTCTAAATACATTCAAATATG 2547
QY 3673 tatccgctcatgagacaataaacccctgataaagtcttcaataataattgaaaaaggaaggt 3732
|||||
Db 2548 TATCCGCTCATGAGACAATAAACCCCTGATAAATGCTTCAATAATATTGANAAGAGAGT 2607
QY 3733 atgagtattcaacatttccgctgctgcgccttattcccttttttgcggcattttgccttcc 3792
|||||
Db 2608 ATGAGTATTCAACATTTCCGCTGCGCCCTTATTCCCTTTTTCGGGCATTTTGCCTTCCT 2667
QY 3793 gttttgctcaccgaaacgcgtggtgaagtaaaagatgctgaagatcagttgggtgca 3852
|||||
Db 2668 GTTTTGTCTCACCACCAAGAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 2727
QY 3853 cgagtgggttacatcgaactggatctcaacagcggtaagatcccttgagtttttcgcccc 3912
|||||
Db 2728 CGAGTGGGTTTACATCGAAGTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC 2787
QY 3913 gaagaaagcttttccaatgatgagcacttttaagttctgctatgtggcgcggtattatcc 3972
|||||
Db 2788 GAAGAAGCTTTTCCAAATGATGAGCAGCTTTTAAAGTTCTGCTATGTGGCGGGTATTATCC 2847
QY 3973 cgtatgacgcggcggaagagcaactcgtcgccgcatacacattctcagaatgacttg 4032
|||||
Db 2848 CGTATTGACGCCCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTG 2907
QY 4033 gttgagtactcaccagtcacagaaaagcatcttacggatggcgatggcgatgaagaatta 4092
|||||
Db 2908 GTTGAGTACTCACAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTA 2967
QY 4093 tgcagtgcctccataaccatgagtgataacactgcggccaaacttacttctgacaaagatc 4152
|||||
Db 2968 TGCAGTGTCTGCCATTAACCATGAGTGATTAACACTCGCGCCAACCTTACTTCTGACAAACG 3027
QY 4153 ggagggaccgaagagctaaaccgctttttgcacaacatggggaatcatgtaactcgcctt 4212
|||||
Db 3028 GGAGGACCGAAGGAGCTAAACCGCTTTTTCACAAACATGGGGGATCATGTAACTTCGCCCT 3087
QY 4213 gatcgttgggaacgcgagctgaatgaagccatacccaaacgacgagcgtgacacacgagt 4272
|||||
Db 3088 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACACGACGAGCGTGACACCACGATG 3147
QY 4273 cctgtagcaatgacaaacagtttgcgcaaaacttaactggcgaaactacttactctagct 4332
|||||
Db 3148 CCTGTAGCAATGCAACAACTTTCGCAAACTATTAACTGGCGAACTTACTTCTACTAGCT 3207
QY 4333 tcccggcacaacataatagactggagcgagataaagtgcagagaccactctgcgc 4392
|||||
Db 3208 FCCGGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTTCAGAGGACCACTTCTCGCG 3267
QY 4393 tcggcccttcggctggctggtttattgctgataaaactgagccggtgagcggtgggtct 4452
|||||
Db 3268 TCGGCCCTTCCTCGGCTGGCTGTTTATTGCTGATAAACTCTGGAGCGGTGAGCGTGGGTCT 3327
QY 4453 cgcgggtatcaattgcagcactcggggcgagatggtaagccctcccgctatcgtagttatctac 4512
```



```

Db 3328 CGCGTATCATTTGACGACTGGGCCAGATGGTAAGCCCTCCGTATCGTAGTTATCTAC 3387
Qy 4513 acgacggggagtcaggcaactatgatgaacgaataagacagatcgctgagatgggtgcc 4572
Db 3388 ACGACGGGGAGTCAGGCAACTATGATCAACGAATAAGACAGATCGCTGAGATAGTGCC 3447
Qy 4573 tcaactgattaagcattggttaactgacagcaacaagtttactcatatatactttagattgat 4632
Db 3448 TCACGTGATTAAGCATTTGGTAAGTCTACGACCAAGTTTACTCATATATATCTTAGATTGAT 3507
Qy 4633 ttaaaacttcatttttaataaaggatctagtggaagatccctttttgataaactcatg 4692
Db 3508 TTAACAACCTCATTTTAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTAAATCTCATG 3567
Qy 4693 accaaaaaccetttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaaagatc 4752
Db 3568 ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATC 3627
Qy 4753 aaaggatctcttgagatccctttttctgtcgcgtaactcgtgctgttcgaacaaaaaaa 4812
Db 3628 AAAGGATCTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGCATCAACAAAAA 3687
Qy 4813 ccaccgctaccagcggtgtttgttgcgagatcaagagctaccaaactctttttccgaag 4872
Db 3688 CCACCGCTACCAGCGGTGTTGTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAG 3747
Qy 4873 gtaactgcttcagcagcagcagatataccaaatactgctctctcagtagtaccgtagtta 4932
Db 3748 GTAACGTGCTTCAGCAGACGCGCAGATACCAATACTGCTCTTCTAGTAGCGGTAGTTA 3807
Qy 4933 ggcacacacttcaagaactctgtagcagcgcctacatacctcgtctcgtatcctgttta 4992
Db 3808 GCCACCACTTCAAGAACTCTGTAGACCGCGCTACATACCTCGCTCTGCTAATCTGTTA 3867
Qy 4993 ccagtggtcgtcgccagtgycgataaagtcgtgtcttaccgggttgagctcaagagcagtag 5052
Db 3868 CCAGTGGCTGCTCCAGTGGCGATAAAGTCGTCTTACC GGTTGGACTCAAGACGATAG 3927
Qy 5053 ttaccggataaggcgcagcgcgttcggcgtgaacgggggttgcgtgcacacagcccgcttg 5112
Db 3928 TTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAGCCACGCTTG 3987
Qy 5113 gagcgaacgacctacacgcaactgagatatacctacagcgtgagctagagaagcgccagc 5172
Db 3988 GAGCAACGACCTACACCGCACTGAGATACCTTACAGCTGAGCTATGAGAAGCGCCACG 4047
Qy 5173 ctccccgaaggagaaagcgacaggtatccggttaagcgcgaggttcggaacagagagag 5232
Db 4048 CTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGAGAG 4107
Qy 5233 ggcacgagggagcttccagggggaacacgctgggtatcctttagtctcgtcggttttcgc 5292
Db 4108 CGCAGAGGAGCTTCCAGGGGGAACGCTGGTATCTTTATAGTCTCTCGGGTTTCGC 4167
Qy 5293 cactctgaactgagcgtcgatttttgatgactcgtcagggggcgagcctatggaaa 5352
Db 4168 CACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAA 4227
Qy 5353 aacgcagcaacgcgcgccttttaccggttctcgtgcttctgtggtccttttgcacatg 5412
Db 4228 AACGCCAGACGGCGCTTTTACGGTTCTGTGGCTTTTGTGCGCTTTTGTCTCACATG 4287
Qy 5413 ttcttctcgtggttatccccctgattctgtggataaccgctattaccgccttttgagtagct 5472
Db 4288 TTCTTTCTGCTGCTTATCCCTCATTTCTGTGGATAACCGCTATTACCGCCTTTTGTGAGT 4347
Qy 5473 gataccgctcgcgcagcgcgaacgacgcagcgcagcagcagcagcagcagcagcagcagc 5532
Db 4348 GATACCGCTCGCGGACGCGAAGACGACGACGACGACGACGACGACGACGACGACGACG 4407
Qy 5533 gagcgcccaatcacgcaaacgcgcctctccccgcgcgttggcggtattcattatgcagctgg 5592

```

```

Db 4408 GAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGCTTGGCCGATTCATTAAATGCAAGTGG 4467
Qy 5593 cagcagaggtttcccgactggaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 5652
Db 4468 CACGACAGGTTTCCCGACTGGAAAGCGGCGAGTGAAGCGCAAGCAATTAATGTGAGTTAG 4527
Qy 5653 ctcaactattaggaaccccgaggttttacactttacactttatgcttcgagcgtcgtatgtgtgga 5712
Db 4528 CTCACCTAATTAGGCAACCCAGGCTTTTACACTTTATGCTTCCGGCTCGGTATGTTGTGGA 4587
Qy 5713 atttgagcgaggaatacaatttcacacaggaacacgactatgaccatgattcacacca 5767
Db 4588 ATTGTGCGGGATAACAATTTTCACACAGGAACAGCTATGACCAATGATTACGCCA 4642

RESULT 12
E49323/c 3796 bp DNA linear PAT 31-JAN-2002
LOCUS E49323 Infectious cDNA clone of North American porcine reproductive and
DEFINITION respiratory syndrome (PRRS) virus and use thereof.
ACCESSION E49323
VERSION E49323.1 GI:18628054
KEYWORDS JP 2000189178-A/32.
SOURCE unidentifed plasmid.
ORGANISM unidentifed plasmid
plasmids.
REFERENCE 1 (bases 1 to 3796)
AUTHORS Calvert,J.G., George,M. and Welshu,S.H.
TITLE Infectious cDNA clone of North American porcine reproductive and
JOURNAL respiratory syndrome (PRRS) virus and use thereof
PATENT: JP 2000189178-A 32 11-JUL-2000;
PFIZER PROD INC
COMMENT
OS Plasmid
PN JP 2000189178-A/32
PD 11-JUL-2000
PF 21-DEC-1999 JP 1999362186
PR 22-DEC-1998 US 60/113345
PI J GUREGORI CALVERT,MICHAEL GEORGE, SHAKUN HWANG WELSHU PC
C12N1/09,A61K39/12,A61K48/00,A61P31/12,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12N15/00,C12N5/00
CC
FH Key Location/Qualifiers
FT source 1..3796 /organism='plasmid'.
FT Location/Qualifiers
FEATURES
source 1..3796
/organism='unidentifed plasmid'
/db_xref='taxon:45202'
BASE COUNT 954 a 932 c 954 g 956 t
ORIGIN
Query Match 45.6%; Score 2630; DB 6; Length 3796;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3138 actgcccgcgtcttttacacgctcgtgactgggaaacccctgcgttaccaccaactaatcg 3197
Db 3796 ACTGCCGCTCGTTTACAACTGCTGACTGGGAAACCCCTGCGTTACCCAACTTAATCG 3737
Qy 3198 cettgcagcacatcccccttttcgcagctggtgtaataagcgaagggccgcagcagatcg 3257
Db 3736 CTTTCAGCACATCCCTTTTCGCCAGCTGGCGTAAATAGCGAAGAGGCCCGCACCGATCG 3677
Qy 3258 ccttcccaaacagttgcgcagcctgaatggcaatggcgcctgagcgtattttctctc 3317
Db 3676 CCTTCCCAACAGTTGCGCAGCCTGAATGGCAATGGCCCTGATGTCGGGTATTTCTCTCT 3617
Qy 3318 tacgcatctgctggttatttcacacgcgcataatgggtgacctctcagtaaatctgctcga 3377
Db 3616 TACGCATCTGTCGGGTATTTACACCGGCATATGGTGCACTCTCAGTACAACTTGTCTCTGA 3557

```

QY 3378 tgcgcataagtttaagcagccccgcagacacccgcacacccgcctgacgcgccctgacgggc 3437
|||||
Db 3556 tgcgcataagtttaagcagccccgcagacacccgcacacccgcctgacgcgccctgacgggc 3497
|||||
QY 3438 tggctcgtcccgccatccgcttacagacaagctgtaacgctcccgagctgcgatgg 3497
|||||
Db 3496 ttgtctgctcccgccatccgcttacagacaagctgtaacgctcccgagctgcgatgg 3437
|||||
QY 3498 tcagaggttttcacccgtcatcccccgaacgcgcgagacgaaagggcctcgtgatacgcct 3557
|||||
Db 3436 tcagaggttttcacccgtcatcccccgaacgcgcgagacgaaagggcctcgtgatacgcct 3377
|||||
QY 3558 atttttataggttaatgctcatgataataatggttttcttagaagtcagggtggcactttcg 3617
|||||
Db 3376 atttttataggttaatgctcatgataataatggttttcttagaagtcagggtggcacttttcg 3317
|||||
QY 3618 gggaatgctgcgggaacccctattgtttatttttcttaataacattcaaatatgtatcc 3677
|||||
Db 3316 gggaatgctgcgggaacccctattgtttatttttcttaataacattcaaatatgtatcc 3257
|||||
QY 3678 gctcatgagacaataaccctgataaattgcttcaataataattgaaaaaggaaagatgag 3737
|||||
Db 3256 gctcatgagacaataaccctgataaattgcttcaataataattgaaaaaggaaagatgag 3197
|||||
QY 3738 tattcaaatcttcgctgctgcgccttatcccttttttgcggaaatttgccttctctttt 3797
|||||
Db 3196 tattcaaatcttcgctgctgcgccttatcccttttttgcggaaatttgccttctctttt 3137
|||||
QY 3798 tgcctaccagaaacgctggtgaaagttaaaagctgctgaagctcagttggtgacagat 3857
|||||
Db 3136 tgcctaccagaaacgctggtgaaagttaaaagctgctgaagctcagttggtgacagat 3077
|||||
QY 3858 gggttacatcgaaactggaattcaacagcggtaagataccttgaggttttgcgccccgaaga 3917
|||||
Db 3076 gggttacatcgaaactggaattcaacagcggtaagataccttgaggttttgcgccccgaaga 3017
|||||
QY 3918 aqgttttcaaatgatagagcacttttaaaagtctgtatgtgctgagcggttattatccgctat 3977
|||||
Db 3016 aqgttttcaaatgatagagcacttttaaaagtctgtatgtgctgagcggttattatccgctat 2957
|||||
QY 3978 tgacgcggggcgaagcgaactcgtgcgcgcataacactattctcagaatgacttggtga 4037
|||||
Db 2956 tgacgcggggcgaagcgaactcgtgcgcgcataacactattctcagaatgacttggtga 2897
|||||
QY 4038 gtactaccagtcacagaaagcaattctacgataggctgacagtaagaagaattatgcag 4097
|||||
Db 2896 gtactaccagtcacagaaagcaattctacgataggctgacagtaagaagaattatgcag 2837
|||||
QY 4098 tgcgcataaacatgataagataaactgcgcgaactgcggaacttacttctgaaacagatcgagg 4157
|||||
Db 2836 tgcgcataaacatgataagataaactgcgcgaactgcggaacttacttctgaaacagatcgagg 2777
|||||
QY 4158 accgaaggagctaaacgcctttttgcacaacatgggggatcatgtaactcgcttgatcg 4217
|||||
Db 2776 accgaaggagctaaacgcctttttgcacaacatgggggatcatgtaactcgcttgatcg 2717
|||||
QY 4218 ttgggaacggagctgaaatgaagcaataccaaacgacgagcgtgacacacagatgcctgt 4277
|||||
Db 2716 ttgggaacggagctgaaatgaagcaataccaaacgacgagcgtgacacacagatgcctgt 2657
|||||
QY 4278 agcaatggcaacaacgttgcgaactatctaactggcggaacttacttctagcttcccg 4337
|||||
Db 2656 agcaatggcaacaacgttgcgaactatctaactggcggaacttacttctagcttcccg 2597
|||||
QY 4338 gcaacaattaaatagctggtgagcggaataaagtgtgagaccacttctgcgctcgcg 4397
|||||
Db 2596 gcaacaattaaatagctggtgagcggaataaagtgtgagaccacttctgcgctcgcg 2537
|||||
QY 4398 ctttcggcgtggtggtttattgtctgataaaattctggagcggctgagcgtgggtcttcggcg 4457
|||||
Db 2536 ctttcggcgtggtggtttattgtctgataaaattctggagcggctgagcgtgggtcttcggcg 2477
|||||
QY 4458 tatcattgacacactgggggccaagattggtaagccctcccgctatcgtagtttatctaacacgac 4517
|||||

Db 2476 tatcatgtgacgacttgggccagatggttaagccctccgcttatgttattcttacacgac 2417
|||||
QY 4518 gggagctcgggcaactatgatatgaaagaaatagacagatacgctgagataggttgcctcact 4577
|||||
Db 2416 gggagctcaggcaactatgatatgaaagaaatagacagatacgctgagataggttgcctcact 2357
|||||
QY 4578 gattaagcatttgtaactgttcagaccaaagtttaactcatatatacttttagattgattaa 4637
|||||
Db 2356 gatttaagcatttgtaactgttcagaccaaagtttaactcatatatacttttagattgattaa 2297
|||||
QY 4638 acttcaatttttaattcaaaagatctaggtgaagataccttttttgataatactcatgaccaa 4697
|||||
Db 2296 acttcaatttttttaattcaaaagatctaggtgaagataccttttttgataatactcatgaccaa 2237
|||||
QY 4698 aatcccttaacgtaggttttcttcacacgagcgtcagaccccgtagaagaaagcaag 4757
|||||
Db 2236 aatcccttaacgtaggttttcttcacacgagcgtcagaccccgtagaagaaagcaag 2177
|||||
QY 4758 atcttcttgagatccctttttctgcgcgtaactctgctgcacacacacacacacac 4817
|||||
Db 2176 atcttcttgagatccctttttctgcgcgtaactctgctgcacacacacacacacac 2117
|||||
QY 4818 gctaccagcgggtggtttgtttgctgcggatcaagagctaccacactcttttccgaagtaac 4877
|||||
Db 2116 gctaccagcgggtggtttgtttgctgcggatcaagagctaccacactcttttccgaagtaac 2057
|||||
QY 4878 tggcttcagcagcgcagacataccaaactctcctctctagtgtagccttagtgagcca 4937
|||||
Db 2056 tggcttcagcagcgcagacataccaaactctcctctctagtgtagccttagtgagcca 1997
|||||
QY 4938 ccacttcaagaactctgtgacacgcctacatacctcgcctctgcttaactctgtttaccagt 4997
|||||
Db 1996 ccacttcaagaactctgtgacacgcctacatacctcgcctctgcttaactctgtttaccagt 1937
|||||
QY 4998 ggcctgcgcagtggcgataaagctgctgttaccgggttggaactcaagacgatagttacc 5057
|||||
Db 1936 ggcctgcgcagtggcgataaagctgctgttaccgggttggaactcaagacgatagttacc 1877
|||||
QY 5058 ggaataaggcgcagcgcgtcggggtgaaacgggggttcgctcacacacacacacacac 5117
|||||
Db 1876 ggaataaggcgcagcgcgtcggggtgaaacgggggttcgctcacacacacacacacac 1817
|||||
QY 5118 aacgacctacacggaactgagatacctacagcgtgagctatgagaagcgccacgcctcc 5177
|||||
Db 1816 aacgacctacacggaactgagatacctacagcgtgagctatgagaagcgccacgcctcc 1757
|||||
QY 5178 cgaaggagaagcgcgacaggtatccggttaagcgcgagggctcggaacagggagagcgac 5237
|||||
Db 1756 cgaaggagaagcgcgacaggtatccggttaagcgcgagggctcggaacagggagagcgac 1697
|||||
QY 5238 gagggagcttccagggggaacgcctggtatctttatagtcctcgtcggttttcgcccact 5297
|||||
Db 1696 gagggagcttccagggggaacgcctggtatctttatagtcctcgtcggttttcgcccact 1637
|||||
QY 5298 ctgacttgagcgtcgatatttttbtgatagtcgtcagggggcgagcctatgaaaaacgc 5357
|||||
Db 1636 ctgacttgagcgtcgatatttttbtgatagtcgtcagggggcgagcctatgaaaaacgc 1577
|||||
QY 5358 cagcaacgcggccttttttaacggttcccgcttttgcctggccttttgcatacatgttctt 5417
|||||
Db 1576 cagcaacgcggccttttttaacggttcccgcttttgcctggccttttgcatacatgttctt 1517
|||||
QY 5418 tctcgtgcttaccctgattctgtgataaacggtattacgccttttgcctggagctgatac 5477
|||||
Db 1516 tctcgtgcttaccctgattctgtgataaacggtattacgccttttgcctggagctgatac 1457
|||||
QY 5478 cgctcgcgcagccgaacacccagcgcagcagtcagtgagcgaagcgaagagcg 5537
|||||
Db 1456 cgctcgcgcagccgaacacccagcgcagcagtcagtgagcgaagcgaagagcg 1397
|||||
QY 5538 cccaatacgcgaacacccgccttcccgcggttggccgattcatttaattgagctggcaaga 5597
|||||

```

Db 1396 CCCAATACGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGTCAGCTGGCAGCA 1337
QY 5598 caggttcccgactgaaagcggcagtgagcgcaacgcaatgaatgagtagctcaac 5657
Db 1396 CAGGTTTCCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGCTCAC 1277
QY 5658 tcattaggcaccacaggttttacacttatgtctccggtctgtatgtgtggaattgt 5717
Db 1276 TCATTAGGCACCCAGGCTTACACTTATAGCTTCCGGCTCGTATGTGTGGTAATGT 1217
QY 5718 gaggcgatacaatttcacacaggaagaaacagctatgaccatgattacgcca 5767
Db 1216 GAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 1167

RESULT 13
LOCUS U02454 5452 bp DNA circular SYN 29-MAR-1996
DEFINITION Cloning vector pCMVEBNA, complete sequence.
ACCESSION U02454
VERSION U02454.1 GI:413820
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE Cloning vector pCMVEBNA.
JOURNAL artificial sequence; vectors.
REFERENCE 1 (bases 1 to 5452)
AUTHORS Kitts,P.A.
TITLE Cloning Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5452)
AUTHORS Swirski,R.A., Van Den Berg,D., Murphy,A.J., Lambert,C.M.,
Friedberg,E.C. and Schimke,R.T.
TITLE Improvements in the Epstein-Barr-based shuttle vector system for
JOURNAL direct cloning in human tissue culture cells
REFERENCE 3 (bases 1 to 5452)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
call (415) 424-8222 or (800) 662-2566, extension 1. International
customers, please contact your local distributor. For technical
information, call (415) 424- 8222 or (800) 662-2566, extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH; this vector has not been
completely sequenced. If you suspect there is an error in this
sequence, please contact CLONTECH's Technical Service Department at
(415) 424-8222 or (800) 662-2566, extension 3 or E-mail
TECH@CLONTECH.COM.

FEATURES
source Location/Qualifiers
1 .5452
/organism="Cloning vector pCMVEBNA"
/db_xref="taxon:31797"
BASE COUNT 1362 a 1246 c 1736 g 1108 t
ORIGIN

Query Match 45.6%; Score 2630; DB 12; Length 5452;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3137 cactggcgcgtgtttacaacgctgactgggaaacccctggcgttaccacaacttaac 3196
Db 2823 CACTGGCGCTCCTTTTACAACGCTGCTGACTGGGAAACCCCTGGCGTTACCCAACTTAATC 2882
QY 3197 gctctgagcacatcccccttttcgacgtggcgtaataagcaagagcccgaccgatc 3256
Db 2883 GCTTTCGACACATNCCCTCTTTCGGCAGCTGGCGTAATAGCGAAGAGGCCCGCCACCGATC 2942
QY 3257 gcccttcccaacagttgcgacgctgaatggcggaatggcgctgatgggtattttccc 3316

```

```

Db 2943 GCCCTTCCCAACAGTTGCCGACGCTGAATGGCGAATGGCGCCTGATGCGGTTATTTCTCC 3002
QY 3317 ttacgcattctgctggtattttcacaccgcataatggcgactctcagtaacaactgctctg 3376
Db 3003 TTACGCATCTGTGCGGTATTTTACACACCGCATATGGTGGCACTCTCAGTACAATCTGCTCTG 3062
QY 3377 atgcgcatagtagtaagcagcccgacaccccgcaacaccccgctgacgcgcctgacggg 3436
Db 3063 ATGCGCGCATAGTTAAGCCAGCGCCGACACCGCCCAACACCGCGCTGACGCGCGCTGACGGG 3122
QY 3437 ctgtctgctcccgccatccgcttacagacaagctgtgacctctccggagagctgcat 3496
Db 3123 CTTGTCTCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCGATGT 3182
QY 3497 gtcagaggttttcacgcgtcatcaccgaaacgcgcgagacgaagggcctcgtgatacgc 3556
Db 3183 GTACAGAGGTTTTCACCGTCATCACCGAAAGCGCGAGAGAAAGGCGCTCGTGATACGCC 3242
QY 3557 tattttataggttaatgcatgataataatggtttcttagacgctgagtggaacttttc 3616
Db 3243 TATTTTATAGGTTAATGTGATGATAAATGTTTCTTAGACGTCAGGTGGCACTTTTC 3302
QY 3617 ggggaaatgctgcgcgaacccctattgtttattttttctaaatacattcaaatatgatac 3676
Db 3303 GGGGAAATGTGCGCGAACCCCTATTTGTTTATTTTCTTAAATACATTCAAATATGTATC 3362
QY 3677 cgcctatgagacaataaacctcgtataatgcttcaataatgataaaagaaagagagatga 3736
Db 3363 CCCTCATGAGACAAATACCCCTGATAAATGCTTCAATAATATTGAAAAAGAGAGATATGA 3422
QY 3737 gttatcaacatttcgctgcgccttattcccttttttgcggcattttgccttccctgttt 3796
Db 3423 GTATTCAACATTTCCGCTGTCGCCCTTATTCCTTTTTCGCGCATTTTGGCTTCCCTGTTT 3482
QY 3797 ttgctcaccgcgaacgcgtggaagtaaaagatgctgaaagatcagttgggtgcaagag 3856
Db 3483 TTGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGTCACGAG 3542
QY 3857 tgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcccccgaag 3916
Db 3543 TGGGTTACATCGAATCGGATCTCAACAGCGGTAAAGATCCTTCAGAGTTTTCGCCCGAAG 3602
QY 3917 aacgtttccatgatgagcacttttaagttctgctatgtggcggttatattcccgta 3976
Db 3603 AACGTTTCCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCGCGGTATTATCCGTA 3662
QY 3977 ttgacgcgggcaagacaactcgcgcgcacacacacattctcagaatgactggttg 4036
Db 3663 TTGACCGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTG 3722
QY 4037 agtactcaccagtcacagaaaaagcatcttacggaaggatggaatgaagaattatgca 4096
Db 3723 AGTACTCACCAGTCACAGAAAAGCATCTACGGATGGCATGACATGAAGAAATTAATGA 3782
QY 4097 gtgtgccaataaccatgagtgatacaactgcggccaacttactctgacaacagatcgagg 4156
Db 3783 GTGCTGCGCAATAACCATGAGTGAATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG 3842
QY 4157 gaccgaagagctaaacccgtttttgcacaacatgggggatcatgtaactcgccttgatc 4216
Db 3843 GACCAAGAGAGGTAAACCGCTTTTTCACAAACATGGGGATCATGTAACTCCGCTTGATC 3902
QY 4217 gttgggaacccggagctgaatgaagccataccaaacagcagcgtgacacacagatgcctg 4276
Db 3903 GTTGGGAACCGGAGCTGAATGAAGCCATACCAACACGAGCGGTGACACACGATGCCTG 3962
QY 4277 tagcaatggcaaacagttgcgcaaacatttaactggygaactacttactctagttccc 4336
Db 3963 TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGTTCC 4022
QY 4337 ggcaacaattaatagactgagtgaggcgagataaagtgcagaccactctcgctcg 4396

```

Db 4023 GGCACAAATTAATAGACTGGATGGAGCGGATAAAGTTGACGACCACCTTCTCGGCTCGG 4082
QY 4397 cccctccggctggtgtttatctgctgataaaatctgagccggtgagcgtgggtctctcg 4456
Db 4083 CCCTTCGGCTGCTGCTGTTTATTTGCTGTATAATCTGGAGCCGCTGAGCGTGGTCTCGCG 4142
QY 4457 gtaactattgagcactgggggcccagatgggtaagccctcccgatctcgtagttatctacacga 4516
Db 4143 GTATCATTCGAGCACTGGGCCAGATGGTAAGCCCTCCCGTAGTTATCTACACGA 4202
QY 4517 cggggagtcaggcaactatggaatgaacgaataagacagatcgcgtgagatgagtgccctcac 4576
Db 4203 CCGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATGAGTGGCTCAC 4262
QY 4577 tgattgaagcattggaactgctcagaccaaagtttactcaatatatacttttagattgatttaa 4636
Db 4263 TGATTAAAGCATTCGTTAACTGTACAGCAAGTTTACTCATATATACATTAGATTGATTAA 4322
QY 4637 aacttcatttttaattaaaggatctaggtgaagatcccttttttgataatctcatgacca 4696
Db 4323 AACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCCTTTTGATATATCTCATGACCA 4382
QY 4697 aaatcccttaactgagtttctccactgagcgtcagaccctgtagaagaatcaaaag 4756
Db 4383 AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCGCTAGAAAAGATCAAG 4442
QY 4757 gatctcttgagatcctttttctcgcgctaatctgctctgcaacaaacaaacacac 4816
Db 4443 GATCTCTTGAGATCCTTTTCTCGCGCTAATCTGCTGCTGCAACAAAAAACCAC 4502
QY 4817 cgtaccagcgggtggtgtttgttcgggatacgaagatcacaactctttttccgaaggtaa 4876
Db 4503 CGCTACAGCGGTGTTGTTGTCGGGATCAAGAGCTACCAACTCTTTTTCGAAAGTAA 4562
QY 4877 ctggtctcagacagcagacagatacacaataactgtctcttagtgtagcgttagtgacc 4936
Db 4563 CTGGCTTCACAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGTAGGCC 4622
QY 4937 accacttcaagaactctgtagcaccgctcacatcctcgtctgctctgctaatcctgttaccag 4996
Db 4623 ACCACTCAAGNACTCTGTAGCACCCTACATACCTCGCTCTGCTGCTGCTGTTACCAG 4682
QY 4997 tggctgctgcagtgcgataaagtctgtcttaccgggttggactcaagacgtagttac 5056
Db 4683 TGCTGTCTGCAGTGGCGATAAGTCTGTCTTACCGGGTTGGACTCAAGACGATAGTTAC 4742
QY 5057 cgtataagcgcagcgtcgggtgaacgggggttcgtgcacacagccagcgttaggc 5116
Db 4743 CGGATAAGCGCGACGCGTCCGGCTGAACGGGGGTTTCGTGCACACAGCCCGCTTGAGC 4802
QY 5117 gaacgaactacacgaactgagatacctacagcgtgagctatgagaaagcgccacgcttc 5176
Db 4803 GAAGCACTACACGAACCTAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC 4862
QY 5177 ccgaaggagaaaggcggacaggtatcccggtgaagcggcggttcggaacagggagcgca 5236
Db 4863 CCGAAGGGGAGAAAGCGGACAGGTATCCGTTAAGCGCGAGGGTCGGAACAGGAGAGCGCA 4922
QY 5237 cgaggagcttccagggggaaacgctgtatcttttatagctcgtcggttttcgcacc 5296
Db 4923 CGAGGGAGCTTCAGGGGGAACCGCTGTGTATCTTTATAGTCTGTCGGGTTCGCGCAC 4982
QY 5297 tctgacttgagcgtcgaattttttgtagtctgctcagggggcgagcgtatgaaaaacg 5356
Db 4983 TCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCATGGAANAACG 5042
QY 5357 ccagcaacgcggtcttttacggttccctggccttttgcgtggccttttgcacatgttct 5416
Db 5043 CCAGCAACGCGGCTTTTACGGTTCCTGGCCCTTTTGTGCTGGCCCTTTTGTCTACATGTTCT 5102
QY 5417 ttcctgcgttatccctgattctgtgataaacggtattaccgctttgagtgactgata 5476
Db 5103 TTCTCGGTTATCCCTGATCTGTGGATAACCGTATTAACCGCTTTGAGTGAGCTGATA 5162

QY 5477 ccgctcgccgagcccaacgagccgagcgcagcagtcagtgagcaggaagcggaagc 5536
Db 5163 CCCTCGCGCAGCCGACGACGAGCGAGCGAGTCAGTCAGTGAGGAGGAGCGAGAGC 5222
QY 5537 gcccaatacgaacaaacgcctctcccgcgcggttggcgcattcattaatgagctggcaag 5596
Db 5223 GCCCAATACCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGACAGCTGGCAGC 5282
QY 5597 acagggttcccgactgaaagcgagcgagtcagcgaacgcaataatgtgagttagctca 5656
Db 5283 ACAGGTTTCCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGCTCA 5342
QY 5657 ctattagcgcaccccggtttcacactttatgcttccggtcgtatgttgtggaattg 5716
Db 5343 CTCATTAGGCACCCAGGCTTTACACTTTTATGCTTCGGGCTCGTATGTTGTGGAATTG 5402
QY 5717 tgagcggaatacaatttcacacaggaacgctatgacctgattacgcc 5766
Db 5403 TGAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCC 5452

RESULT 14
XXU02433/C
LOCUS
DEFINITION
Cloning vector pNASSbeta, complete sequence.
ACCESSION
U02433
VERSION
U02433.1 GI:413799
KEYWORDS
SOURCE
ORGANISM
Cloning vector pNASSbeta.
Cloning vector pNASSbeta
artificial sequence; vectors.
REFERENCE
1 (bases 1 to 6541)
MacGregor, G. R. and Caskey, C. T.
AUTHORS
Construction of plasmids that express E. coli beta-galactosidase in
TITLE
mammalian cells
JOURNAL
Nucleic Acids Res. 17 (6), 2365 (1989)
MEDLINE
89202057
REFERENCE
2 (bases 1 to 6541)
Kitts, P. A.
AUTHORS
ClONTECH Vectors On Disc version 1.3
TITLE
ClONTECH Vectors On Disc version 1.3
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 6541)
Kitts, P. A.
AUTHORS
Direct Submission
TITLE
Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
JOURNAL
1020 East Meadow Circle, Palo Alto, CA 94303 USA
COMMENT
This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
call (415) 424-8222 or (800) 662-2566, extension 1. International
customers, please contact your local distributor. For technical
information, call (415) 424-8222 or (800) 662-2566, extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH; this vector has not been
completely sequenced. If you suspect there is an error in this
sequence, please contact CLONTECH's Technical Service Department at
(415) 424-8222 or (800) 662-2566, extension 3 or E-mail
TECH@CLONTECH.COM.

FEATURES
source
1. 6541
Location/Qualifiers
/organism="Cloning vector pNASSbeta"
/db_xref="taxon:31024"
BASE COUNT 1549 a 1675 c 1744 g 1573 t
ORIGIN

Query Match 45.6%; Score 2630; DB 12; Length 6541;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3138 actggcgctgttttacaacgtcgtgactgggaaacccctggcgttaccaccaacttaatcg 3197
|||||

Db 6541 ACTGGCCGTCGTTTACAAGTCGCTGACTGGGAAACCCCTGGCGTTACCCAACTTAAATCG 6482
Qy 3198 cettgcagacatcccccttttcgcagctgctgtaataagcgaagggcccgccagctgcg 3257
Db 6481 CTTTCAGACATCCCTTTTCGCAGCTGGCGTAAATAGCGAAGAGGCCGACCGCATCG 6422
Qy 3258 ccttcccaacagttgcgcagcctgaatggcgaatggcgcctgagctgagctatttctcct 3317
Db 6421 CCCTTCCCAACAGTTGCCAGCTGAAATGGGNAATGGGCCCTGATGGCGGTATTTCTCCT 6362
Qy 3318 tacgcactgtgcggtatttcacaccgcataatggcgactctcagtaacaatctgctcga 3377
Db 6361 TACGCATCTGCGGTATTTACACCGCATATGTCGACTCTCAGTACAATCTGCTCTGA 6302
Qy 3378 tgcgcagatagtaagccagccgcagacccgcaccccaacaccgcctgagcgcctgacggc 3437
Db 6301 TGCCGCATAGTTAAAGCCAGCCGACACCCGCCAACAACCCGCTGACGCCCTGACGGC 6242
Qy 3438 ttgtctgctccggcgcctctacagacaagctgtgacgctcctccggagctgcatgtg 3497
Db 6241 TTGTCTGCTCCGGCATCCGCTTACAGACAGCTGTGACCGTCTCCGGAGCTGCTATGTG 6182
Qy 3498 tcagaggtttccacgctcatcccgaaacgcgcgcagagaaaggccctcgtgatacgcct 3557
Db 6181 TCAGAGGTTTTCACCGTCTCATACCCGAAACGCCGAGACGAAAGGCCCTCGTGATACGCT 6122
Qy 3558 attttataggttaatgcatgataataatggtttcttagacgtcaggtgcaacttttcg 3617
Db 6121 ATTTTATAGGTTAATGTCTATGATTAATGGTTTCTTAGACGTGAGGTGGCACTTTTCG 6062
Qy 3618 gggaaatgcgcggaacccctatttcttttctaaatacatcattcaaatatgtatcc 3677
Db 6061 GGGAAATGTGGCGGNAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGTATCC 6002
Qy 3678 gctcatgagacaataaccctgataaattgcttcaataatattgaaagaagagatgatgag 3737
Db 6001 GCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAGAGAGATGAG 5942
Qy 3738 tattcaaacatttcgctgctgcgccttatccctttttctggcgcatcttgccctccgtttt 3797
Db 5941 TATTCAACATTTCCGTGTCGGCCCTTATTCCCTTTTTCGGCATTTTTCCTTCCCTGTTT 5882
Qy 3798 tgcaccaccgaagaacgctgggtgaagtaaaagatgctgaagatcagttgggtgacagagt 3857
Db 5881 TGCTCACCCAGNAAGCTGTGGTGAAGTAAAGATGCTGNAGATCAGTTGGGTGCACAGAT 5822
Qy 3858 gggttacatcgaactggatctcaacagcggtaagatccttgagagtttccgcccgaaga 3917
Db 5821 GGGTTACATCGAATCTGATCTCAACAGCGGTAAAGATCCTTGAGAGCTTTTCGCCCGAAGA 5762
Qy 3918 acgttttccaatgatgagcaactttttaaagtctgctatgtggcggttatataccgctat 3977
Db 5761 ACGTTTTCCATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATTATCCCGTAT 5702
Qy 3978 tgaecgggggaagcaactcggctgcgcgcatacactattctcagaatgacttgggttga 4037
Db 5701 TGACCGCGGCAAGAGCAACTCGGTTCGGCATACACTATTCTCAGATGACTTGGTTGA 5642
Qy 4038 gtactcaaccagtacagaaaaagcatcttaccgtagtgccatgagtaagaaattatgcag 4097
Db 5641 GTACTCACCACTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATATGCAG 5582
Qy 4098 tgcgtcctaaccatgagtgataaacacttcggcgcaacttacttctgcacacagatcggagg 4157
Db 5581 TGCTGCCAATAACCATGATGTGATAACACTTCGGGCCAACTTACTTCTGACAACGATCGGAGG 5522
Qy 4158 accgaaggagctaaacgcttttttcacacaactggggatcctgtaactcgccttgatcg 4217
Db 5521 ACCGAAGNGCTAACCGCTTTTTCACACACATGGGGGATCATGTAACCTCGCTTGATCG 5462
Qy 4218 ttgggaacccgagctgaatgaagccataccaaacacagagcgtgacacacagatgcctgt 4277
Db 5461 TTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACACGATCCGCTG 5402

Qy 4278 agcaatggcaacaaacgcttgcgcacaaactatttaactggcgaactacttactctagctcccg 4337
Db 5401 AGCAATGGCAACACAGCTTTCGCAAACTATTAACTGCGCAACTACTTACTTACTAGTTCGCG 5342
Qy 4338 ccaacaataatagactgagtgagcgagataaagtgtcagagaccacttctcgtcgcgcg 4397
Db 5341 GCACAATAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACCTTCTCGGCTCGCG 5282
Qy 4398 ccttccgctggctgttatttgcgtgataaactcgtgagccggtgagcggtggtctcgcg 4457
Db 5281 CTTTCGGCTGGCTGTTTATTTGCTGATAAATCTGAGCGCGTGAAGGCTGCTCGCG 5222
Qy 4458 tatcattgcagcaactcggggccagatggtaagccctcccgatcgtagttatctacacgac 4517
Db 5221 TATCATTCAGCACTGGGCGCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC 5162
Qy 4518 ggggagtcaggaacatctggtatgagcaaatagacagatccgctgagatgagctccact 4577
Db 5161 GGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCACT 5102
Qy 4578 gattaagcattggtaactgtcagaccacagtttactcatatatactatttagatgtatttaa 4637
Db 5101 GATTAAAGCATTTGCTAACCTGCAGACCAAGTTTACTCATATATACATTTAGATTGATTTAA 5042
Qy 4638 acttcaatttttaatttaaaagatctagtggaagatccctttttgataatctcatgacaaa 4697
Db 5041 ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTATTAATCTCATGACCAA 4982
Qy 4698 aatcccttaacgtgagtttctccactgagcgtcagaccgcgtagaaaaagatcaagg 4757
Db 4981 AATCCTTAACTGAGTTTTCCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGG 4922
Qy 4758 atctcttgagatccctttttctgcggtaatctgctgtgcaacaaaaaacccacc 4817
Db 4921 ATCTTCTTGAGATCCTTTTTCGCGGTAACTGCTGCTTGCAACAAAAAACACCACC 4862
Qy 4818 gctacacgcggtgggttctgttcgcggaataagagctacaaactcttttccgaaggtaac 4877
Db 4861 GCTACACGCGGTGGTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAAC 4802
Qy 4878 tggcttcagcagagcgcagataccaaatactctctcttagtgtagcgttagtgagca 4937
Db 4801 TGGCTTACAGAGCGCGAGATACCAAAATACTGCTCCTCTAGTGTAGCCGTAGTTAGGCCA 4742
Qy 4938 ccaactcaagaactctgtagcaccgcctacatacctcgtctcgtctgtaactctgtaccagt 4997
Db 4741 CCACCTTCAAGAACTCTGTAGCACCCGCCCTAGATACCTCGCTCTGCTAACTCTGTACCAGT 4682
Qy 4998 ggtcgtgccagtgcgataagctcgtcttaccgggttggactcaagacgtagttacc 5057
Db 4681 GGCTGCTGCCAGTGGCGATAAGTCTGTCTTTACC GGTTGGACTCAAGACGATAGTTACC 4622
Qy 5058 ggaataaggcgcgctggtgctgacgggggttcgtgcacacagccagcttggagcg 5117
Db 4621 GGATGAAGCGCAGCGGTGGGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGGAGCG 4562
Qy 5118 aacgacctacaccgaactgagatacctacagcgtgagctatgagaaaagccacacttcc 5177
Db 4561 AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATCAGAAAAGGCCACGCTTCC 4502
Qy 5178 cgaagggaagagcgagcaggtatcccgtaagcggcgaggttcggaacagggagcgac 5237
Db 4501 CGAAGGGAGAAGGGCGACAGGTATCCGGTAAGCGCGAGGGTCTGGAAACAGGAGCGCAC 4442
Qy 5238 ggggagcttcaggggaaacgctgtatctttagtctcgtggttttcgcaact 5297
Db 4441 GAGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCTCTGCGGGTTTCGCCACCT 4382
Qy 5298 ctgacttgagcgtcgtattttgtgctcgtcagggggcgagcgtatggaacacgc 5357
Db 4381 CTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGAANAACGC 4322

Qy	5358	cagcaacgcgcgccttttaacggttcctggcccttttgctggcccttttgagtcacatggtttt	5417
Db	4321	CAGCAACGCGGCCCTTTTACGGTTCTTGGCCCTTTTGGCTGGCTTTTGGCTCACATGTTCTT	4262
Qy	5418	tctcggtatcccttgattctggtgaataacccgttatccgcctttgagtcagctgatac	5477
Db	4261	TCCTGCGTTATCCCTGTATCTGTGGATTAACCGTATTACCGCCCTTGTAGTGAGCTGATAC	4202
Qy	5478	cgctcgccgacgcgcaaacgacgcgagcgcagcgagtcagtgagcgcgagaacgcggaagagcg	5537
Db	4201	CGCTCGCGCAGCGCAACGACGAGCGCAGCGAGTCACTGAGCGAGGAAAGCGAAGAGCG	4142
Qy	5538	ccaatacgcgaacacgcgcctctcccgcgctgttgccgattcattaaatgcagctggcacga	5597
Db	4141	CCCAATACGCAAAACCGGCTCTCCCGCGCGTGGCCGATTATTAAATGACCTGGCACGA	4082
Qy	5598	cagcttcccgactgaaacgcgcgcagtcgagcgcacgcgcaataatgtgagctagctcac	5657
Db	4081	CAGTTCCTCCCGACTGGAAACGCGCGCAGTCAGCGCAACGCCAATTAATGTGAGTTAGCTCAC	4022
Qy	5658	tcattaggcaccccgagctttacactttatgcttcccgctcgatgttgttgaggaaattgt	5717
Db	4021	TCATTAGGCACCCCGAGGCTTTACACTTTATGCTTCGGCTCGTATGTTGTGTGCAATTGT	3962
Qy	5718	gagcggatacaaatctcacacaggaacagctatgaccatgattacgcga	5767
Db	3961	GAGCGGATACAAATTTACACACAGGAAACAGCTATGACCATGATTACGCCA	3912
RESULT 15			
XXU02435/c			
LOCUS	XXU02435	5894 bp	DNA
DEFINITION	Cloning vector pSVbeta, complete sequence.		
ACCESSION	U02435	circular SYN 29-JAN-1997	
VERSION	U02435.1	GI:413801	
KEYWORDS	Cloning vector pSVbeta.		
SOURCE	Cloning vector pSVbeta		
ORGANISM	artificial sequence; vectors.		
REFERENCE	1 (bases 1 to 6894)		
AUTHORS	MacGregor,G.R. and Caskey,C.T.		
TITLE	Construction of plasmids that express E. coli beta-galactosidase in mammalian cells		
JOURNAL	Nucleic Acids Res. 17 (6), 2365 (1989)		
MEDLINE	89202057		
REFERENCE	2 (bases 1 to 6894)		
AUTHORS	Kitts,P.A.		
TITLE	CLONTECH Vectors On Disc version 1.3		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 6894)		
AUTHORS	Kitts,P.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-OCT-1993); Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA		
COMMENT	This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424- 8222 or (800) 662-2566, extension 3. This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH; this vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.		
FEATURES	Location/Qualifiers		
SOURCE	1.. 6894		
	/organism="Cloning vector pSVbeta"		
	/db_xref="taxon:31847"		
BASE COUNT	1630 a 1780 c 1830 g 1654 t		
ORIGIN			

Qy	4158	accgaaggagctaacccgctttttgcacaacatgggggatcatgtactatgcgcttgatcg	4217
Db	5874	ACGGAAGGAGCTAACGCGCTTTTTCACAAACATGGGGGATCATGTAACTGCGCTTGATCG	5815
Qy	4218	ttgggaaccggagctgaatgaagaccataccaagaacagagcgtgaacaccagatgcctgt	4277
Db	5814	TTGGGAACCGGAGCTGATGAAGCCATACCAACGACGAGCGTGACACGATGCCCTGT	5755
Qy	4278	agcaatggcaacaacgcttgcgcaaaacttaacttgcgaaactacttactctagtctcccg	4337
Db	5754	AGCAATGGCAACAACGCTTGCSCAAACTATTAACTTGGCGAACTACTTACTCTAGCTCCCG	5695
Qy	4338	gcaacaattaatagactgtagtgaggcgcggaataaagcttcgaggacaactcttgcgctgcgc	4397
Db	5694	GCAACAATTAATAGACTGGATGGAGCGGATAAAGTTTCGAGGACCACTTCTGCGCTCGC	5635
Qy	4398	ccttcgcgctgcgctggtttatctgctgataaaactctgagaccggtgaagcgtgggtctcgcg	4457
Db	5634	CCTTCCGGCTGGCTGGTTTATTTGCTGTAAATCTGGAGCCGGTGAAGCTTTCGCGG	5575
Qy	4458	tatcattgcagcactggggccagatggtaagccctcccgtatcgttagttactacagcac	4517
Db	5574	TATCATTTGAGCACATGGGCCAGATGGTAAGCCCTCCCGTATCTGTAGTTATCTACACGAC	5515
Qy	4518	ggggagtcaggcaactatggaatgaacgaaatagacagatcgtcgatagataggtgcctcact	4577
Db	5514	GGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTGCTCACT	5455
Qy	4578	gattaagcattgtaactcgacgaagtttactcatatatactttagattgatttaaa	4637
Db	5454	GATTAAAGCATTGTGACTCTGACACAGTTTACATATATATATATATCTGTAGATTGATTTAA	5395
Qy	4638	acttcatttttaatttaaaagatctagtgaaagactcttttggataactctcatgaccaa	4697
Db	5394	ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCAA	5335
Qy	4698	aatccttaacgtgagcttctgtccactgagcgtcagaccgccgtagaagaatcgaag	4757
Db	5334	AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGG	5275
Qy	4758	atcctcttgagatcccttttctgcgagtaactgcgtctgcaaacaaaaaacacc	4817
Db	5274	ATCTTCTTGAGATCCCTTTTTCGGCGGTAACTGCTGTGCTTGCAAAACAAAAAACACC	5215
Qy	4818	gctaccagcgggtggttgtttgcgggatcaagagactaccaactcttttccgaaggtaac	4877
Db	5214	GCTACCAACGGTGTGTGTTTTCGGGATCAGAGCTACCACTCTTTTCCGAAGGTATAC	5155
Qy	4878	tggcttcagcagcgagataccaaaactgtccttctagttagcgttagttaggcca	4937
Db	5154	TGGCTTCAGCAGACGCACATACCAAACTGCTCTTCTAGTGTAGCGTAGTTAGGCCA	5095
Qy	4938	ccacttcaagaactctgtagaccgctacatactgcgtctgtaactcctgttaccagt	4997
Db	5094	CCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGTTAATCTGTGTACCGAGT	5035
Qy	4998	ggctgctgccagtggcgataaagtcgtcttaccgggttggaactcaagacgatatgtacc	5057
Db	5034	GGCTGCTGCCAGTGGCGATAGTCGTGCTTACCGGGTGGCACTCAAGACGATAGTTACC	4975
Qy	5058	ggataaggcgacggtcggggtgaacggggggttcgtgcacacagcccagcttgagcg	5117
Db	4974	GGATAAGGGCGACGGCTCGGGCTGAACGGGGGTTCTGTGCACACAGCCACAGCTTGGAGCG	4915
Qy	5118	aacgacctacaccgaactgagatacctacagcgtgagctatgaagaagcgccagcttcc	5177
Db	4914	AACGACCTTACACCGAACTGAGTATACCTACGCGTGAGTATGAGAAAAGCCGACGCTTCC	4855
Qy	5178	cgaagggaagaaggcgacaggtatccggtaagcggcaggttcggaacaggaagcgcac	5237
Db	4854	CGAAGGGAGAAAGCGGACGACAGGTATCCGGTAAACGGCGAGGGTCTGGNACAGGAGACCGAC	4795
Qy	5238	gaggagacttccaggggggaacgcctggttatctttagtctgctgcgggtttgcgcacct	5297

Db	4794	GAGGAGAGCTCCAGGGGAAACGCCTGGTATCTTTTATAGTCGTGCGGGTTTCGCACCT	4735
Qy	5298	ctgacttgagcgtcgatatttctgatgctcgtcagggggcgagcctcatggaataacgc	5357
Db	4734	CTGACTTGAGCGTCGATTTTGTGATGCTGCTGAGGGGGGAGCCCTATGGAATAACGC	4675
Qy	5358	cagcaaacgaggccttttacggttgccttgcccttttgccttgccctttgctcacatgtctt	5417
Db	4674	CAGCAAACGCGGCCTTTTACGGTTCTTGGCCTTTTGTGCTGGCCTTTTGTCTCACATGTTCTT	4615
Qy	5418	tcttcggttatccctgattctctgtgataaacgftattaccgctcttgagtcgagtcgatac	5477
Db	4614	TGCTCGGTATATCCCTCGATCTCTGCGATAACCGTATTACCGCTTTTGAGTGAGCTGATAC	4555
Qy	5478	cgtctgcgcgagccgaacacgcagcagcgcagcgagtcagtgagcggaggaagcgaagcgc	5537
Db	4554	CGCTCGCGCAGCCGAAACGACCGACGCGAGTCAGTGACCGAGGAAGCGGAACACGG	4495
Qy	5538	cccaatacgcgaacccgcctctcccgcgcgttgccgattcattaatcgagctggcacga	5597
Db	4494	CCCAATAGCGAAACGCGCTCTCCCGCGCTTGCGCCGATTCAATTATGCACTGGCAGCA	4435
Qy	5598	caggtttcccgaactgaaagcgggcagtgagcgcgaacgaattaatgtgagttagctcac	5657
Db	4434	CAGGTTTCCCAGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTATGTGAGTTAGCTCAC	4375
Qy	5658	tcattaggcacccccaggctttacactttatgcttcccgctcgatgtgtgtggaattgt	5717
Db	4374	TCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTGTGTGGGAATTGT	4315
Qy	5718	gagcggataacaatttcacaggaataacgctatgaccatgattacgcga	5767
Db	4314	GAGCGGATTAACAATTTTCACAGGAACAAGCTATGACCATGATTACGCCA	4265

Search completed: August 31, 2002, 22:41:47
Job time: 25536 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:58:22 ; Search time 824.51 Seconds
(without alignments)
3592.045 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgagggcccccagtgctct.....aattgctcagcgccacctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 1736436 seqs, 858457221 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1725	100.0	1725	24	Synthetic human ac
2	1722	99.8	1845	24	Human cDNA encodin
3	1722	99.8	2256	16	Human acetylcholin
4	1712	99.2	5767	24	Plasmid pTM034. S
5	1712	99.2	14446	24	Plasmid pTM036. S
6	1252	72.6	2253	11	Sequence encoding
7	1070	62.0	12113	24	Human acetylcholin
8	489	28.3	1800	11	Sequence encoding
9	35	2.0	35	17	Acetylcholinestera

10	33	1.9	2089	24	ABI99790	Mouse ischaemic co
11	28	1.6	8418	22	AAS41906	Genomic sequence #
12	27	1.6	816	23	AAS74821	DNA encoding novel
13	27	1.6	1532	23	AAS74822	DNA encoding novel
14	27	1.6	1538	22	AAD09548	Human protease pro
15	25	1.4	496	23	AAS74456	DNA encoding novel
16	25	1.4	2523	22	AAF82995	DNA encoding novel
17	25	1.4	4975	23	AAS74457	DNA encoding novel
18	22	1.3	4173	21	AAA62899	Murine JNK3 bindin
19	22	1.3	4200	21	AAA62900	Murine JNK3 bindin
20	22	1.3	4266	21	AAA62902	Murine JNK3 bindin
21	22	1.3	4269	21	AAA62901	Murine JNK3 bindin
22	21	1.2	21	24	AAS17550	Human acetylcholin
23	20	1.2	20	19	AAS17550	Nuclease resistant
24	20	1.2	20	19	AAV54213	Nuclease resistant
25	20	1.2	20	19	AAV54215	Nuclease resistant
26	20	1.2	20	19	AAV41281	Human AChE mRNA sp
27	20	1.2	20	19	AAV41282	Human AChE mRNA sp
28	20	1.2	20	19	AAV41283	Human AChE mRNA sp
29	20	1.2	20	22	AAH44814	Antisense oligonuc
30	20	1.2	20	22	AAH44815	Antisense oligonuc
31	20	1.2	20	22	AAH44817	Antisense oligonuc
32	20	1.2	1698	20	AAH86613	cDNA encoding an a
33	20	1.2	2369	23	ABL13421	Drosophila melanog
34	20	1.2	2478	20	AAH86614	cDNA encoding an a
35	20	1.2	5124	23	AAS95024	DNA encoding human
36	20	1.2	6900	23	AAS95022	DNA encoding murin
37	20	1.2	6901	23	AAS95027	DNA encoding murin
38	20	1.2	7020	24	AAD26312	Human G-protein co
39	20	1.2	7119	23	AAS95023	DNA encoding murin
40	20	1.2	7119	23	AAS95025	DNA encoding murin
41	20	1.2	7172	23	AAS95026	DNA encoding human
42	20	1.2	7694	22	ABA07755	Human ovarian and
43	20	1.2	7694	22	AAL03544	Human reproductive
44	20	1.2	18413	23	ABL13420	Drosophila melanog
45	19	1.1	19	24	AAS17551	Human acetylcholin

ALIGNMENTS

RESULT 1
AAS17549
ID AAS17549 standard; DNA; 1725 BP.
XX
AC AAS17549;
XX
DT 25-FEB-2002 (first entry)
XX
DE Synthetic human acetylcholinesterase gene.
XX
KW Human; acetylcholinesterase; AChE; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.
OS Homo sapiens.
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

Db 1561 aatgagcccgagaccacagggcccccacaaatgccccgtacacgcg9ggggccacagcag 1620
 QY 1621 tacgttagctggaacctgcgcgcgtgaggtgcgcg9ggggtgctgcgcgcacagcctgc 1680
 Db 1621 tacgttagctggaacctgcgcgcgtgaggtgcgcg9ggggtgctgcgcgcacagcctgc 1680
 QY 1681 gcctctggaacgcctctcccccacaaattgctcagcgcacacctga 1725
 Db 1681 gcctctggaacgcctctcccccacaaattgctcagcgcacacctga 1725

RESULT 2

AAS17493
 ID AAS17493 standard; cDNA; 1845 BP.
 AC AAS17493;
 XX
 XX 26-FEB-2002 (first entry)
 XX Human cDNA encoding acetylcholinesterase, AChE.
 DE Human; ss; AChE; acetylcholinesterase; polymorphic variant; haplotyping;
 KW genotyping; neurological disease; Parkinson's disease;
 KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 1..1845
 FT CDS
 FT /*tag= a
 FT /product= "ACHE"
 FT /replace (36,T)
 FT /*tag= b
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (101,A)
 FT /*tag= c
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (183,C)
 FT /*tag= d
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (838,G)
 FT /*tag= e
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (873,T)
 FT /*tag= f
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (1032,A)
 FT /*tag= g
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (1057,A)
 FT /*tag= h
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 FT /replace (1431,T)
 FT /*tag= i
 FT variation
 FT /standard_name= "Single nucleotide polymorphism"
 XX WO200179219-A2.
 XX
 XX 25-OCT-2001.
 XX
 XX 11-APR-2001; 2001WO-US11853.
 XX
 XX 14-APR-2000; 2000US-197173P.
 XX
 XX (GENA-) GENAISSANCE PHARM INC.
 XX (KAZE/) KAZEMI A.
 XX
 XX Bentivegna SC, Chew A, Choi JY, Koshy B;
 XX WPI; 2002-055248/07.
 DR P-PSDB; AAU11231.
 XX

PT New polymorphic variants comprising acetylcholinesterase (ACHE)
 PT isogene, useful in expressing AChE protein for use in screening for
 PT candidate drugs to treat diseases related to AChE activity, e.g.
 XX neurological diseases or cancer -
 PS Claim 26; Fig 2; 79pp; English.
 XX
 CC The invention relates to a polynucleotide comprising a polymorphic
 CC variant of an acetylcholinesterase (ACHE) gene or fragment, protein or
 CC complement, the variant comprising an AChE isogene defined by a haplotype
 CC selected from haplotypes 1-20 listed in the specification. Also included
 CC are methods for haplotyping and genotyping the AChE gene of an
 CC individual, a method for predicting a haplotype pair for the AChE gene of
 CC an individual, a method for identifying an association between a trait
 CC and at least one haplotype or haplotype pair of AChE gene, recombinant
 CC nonhuman organisms transfected or transformed with the polynucleotide
 CC where the organism expresses AChE protein encoded by the first
 CC nucleotide sequence or encoded by the polymorphic variant sequence,
 CC an isolated antibody specific for and immunoreactive with AChE,
 CC a method of screening for drugs targeting the polypeptide contacting AChE
 CC polymorphic variant with a candidate agent and assaying for binding
 CC activity, a computer system for storing and analysing polymorphism data
 CC for AChE gene and a genome anthology for AChE gene which comprises AChE
 CC isogenes defined by haplotypes 1-20 given in the specification.
 CC The polymorphisms are useful for studying the biological function of
 CC AChE as well as in identifying drugs targeting this protein for the
 CC treatment of disorder related to its abnormal expression or function.
 CC The polymorphic variants may also be used in screening for compounds
 CC targeting AChE to treat a specific condition or disease predicted to be
 CC associated with AChE activity e.g. neurological diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
 CC and tumours. The AChE gene maps to human chromosome 7q22. The
 CC present sequence is the coding sequence of the AChE gene.
 XX
 SQ Sequence 1845 BP; 302 A; 590 C; 608 G; 345 T; 0 other;

Query Match 99.8%; Score 1722; DB 24; Length 1845;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 atgagggcccgagctgtctgtcacacgcttccctgcttccctccactctctctc 60
 Db 1 atgagggcccgagctgtctgtcacacgcttccctgcttccctccactctctctc 60
 QY 61 ctctctggctcctgggtgagaggtggtgggctgagggcgaggagtcagagctgctg 120
 Db 61 ctctctggctcctgggtgagaggtggtgggctgagggcgaggagtcagagctgctg 120
 QY 121 gtgacggtgctgggtgggcccgcctgctggggcattcgctgaagaccccggggcccctgtc 180
 Db 121 gtgacggtgctgggtgggcccgcctgctggggcattcgctgaagaccccggggcccctgtc 180
 QY 181 tctgtcttctggcattcccttgcgagccaccatggagaccctgctgcttcttgcca 240
 Db 181 tctgtcttctggcattcccttgcgagccaccatggagaccctgctgcttcttgcca 240
 QY 241 ccggagcccaagcagcttggtaggggtggtgtagcgtacacaccttcacagtgctgc 300
 Db 241 ccggagcccaagcagcttggtaggggtggtgtagcgtacacaccttcacagtgctgc 300
 QY 301 taccaatatgtggacacccctataccacaggttttggggcaccgagatgtgaaaccccaac 360
 Db 301 taccaatatgtggacacccctataccacaggttttggggcaccgagatgtgaaaccccaac 360
 QY 361 cgtgagctgagcggagactgccttaacctcaacgtgtggacaccataccccccggcctaca 420
 Db 361 cgtgagctgagcggagactgccttaacctcaacgtgtggacaccataccccccggcctaca 420
 QY 421 tccccacccctgtccctgctgtgctatgtgggggtggtctctacagtgggggcctctccc 480
 Db 421 tccccacccctgtccctgctgtgctatgtgggggtggtctctacagtgggggcctctccc 480

QY 481 ttggacgtgtacagtgagccgcttcttgggtacagggccgagagactgtgtgtgttccatg 540
Db 481 ttggacgtgtacagtgagccgcttcttgggtacagggccgagagactgtgtgtgttccatg 540
QY 541 aactaccgggtggagaccccttggcttccctgcccctgcccggagccgagagcccgccgagc 600
Db 541 aactaccgggtggagaccccttggcttccctgcccctgcccggagccgagagcccgccgagc 600
QY 601 aatgtgggttctcttgatcagaggtggccctgcagtggtgtcagagaaactgtggcagcc 660
Db 601 aatgtgggttctcttgatcagaggtggccctgcagtggtgtcagagaaactgtggcagcc 660
QY 661 ttcgggggtgacccagacatcagtgacgtgttttggggagagcgcgcgcctcctggtg 720
Db 661 ttcgggggtgacccagacatcagtgacgtgttttggggagagcgcgcgcctcctggtg 720
QY 721 ggcagtcacactgtctcccccgcagccgagcccttgccttccacagggccgtctgcagagc 780
Db 721 ggcagtcacactgtctcccccgcagccgagcccttgccttccacagggccgtctgcagagc 780
QY 781 ggtgcccccaatggacccctggccaggtggccatgggagagcccgctgcaggggcccagc 840
Db 781 ggtgcccccaatggacccctggccaggtggccatgggagagcccgctgcaggggcccagc 840
QY 841 cagctggcccacactgtgggctgtctcccaagcgcgcaactgtgtgggaatgacacagagctg 900
Db 841 cagctggcccacactgtgggctgtctcccaagcgcgcaactgtgtgggaatgacacagagctg 900
QY 901 gtagcctgcttcggagacagcagcagcagcagctccctggtgaaccagaaatggcagctgctg 960
Db 901 gtagcctgcttcggagacagcagcagcagcagctccctggtgaaccagaaatggcagctgctg 960
QY 961 cctcaagaagcgtcttcctggttctctctgctgctgtgtgtagtgagacttctcctcagt 1020
Db 961 cctcaagaagcgtcttcctggttctctctgctgctgtgtgtagtgagacttctcctcagt 1020
QY 1021 gacacccagagccctcatcaacgaggagacttccacagcgcctgcaggtgtgtgtggt 1080
Db 1021 gacacccagagccctcatcaacgaggagacttccacagcgcctgcaggtgtgtgtggt 1080
QY 1081 gtggtgaaagtgagggctcgtatttctctggtttacggggcccgagcttcagcaaaagac 1140
Db 1081 gtggtgaaagtgagggctcgtatttctctggtttacggggcccgagcttcagcaaaagac 1140
QY 1141 aacgagtcctcatcagccggccgagttccctggccgggtgtcggtgcgggttcccccag 1200
Db 1141 aacgagtcctcatcagccggccgagttccctggccgggtgtcggtgcgggttcccccag 1200
QY 1201 gtaagtgaactggcagccagagctgtgttctgctgacattacacagactggctgcatccag 1260
Db 1201 gtaagtgaactggcagccagagctgtgttctgctgacattacacagactggctgcatccag 1260
QY 1261 gaccggcagcctgagggagccctgagcgtgtgtgtggcgacacacaaatgtgtgtg 1320
Db 1261 gaccggcagcctgagggagccctgagcgtgtgtgtggcgacacacaaatgtgtgtg 1320
QY 1321 cccgtggccagctggctgggggactggctgccaggggtgcccggtgtctacccctacgtc 1380
Db 1321 cccgtggccagctggctgggggactggctgccaggggtgcccggtgtctacccctacgtc 1380
QY 1381 ttgtgaacacgtgtcttccacgctctctgcccctgtggtatgggggtgcccacacggctac 1440
Db 1381 ttgtgaacacgtgtcttccacgctctctgcccctgtggtatgggggtgcccacacggctac 1440
QY 1441 gtagtcgagttcatctttgggttccccctggaccctctcgaaactacacgagcagagag 1500
Db 1441 gtagtcgagttcatctttgggttccccctggaccctctcgaaactacacgagcagagag 1500
QY 1501 aaaaattctgcccagagactgtatgcatactgtggccaaactttgccgcgacaggggatccc 1560
Db 1501 aaaaattctgcccagagactgtatgcatactgtggccaaactttgccgcgacaggggatccc 1560
QY 1561 aatgagcccgagaccccaaggcccaacaatggcccccggtacacggcggtgtcagcag 1620

Db 1561 aatgagcccgagaccccaaggcccaacaatggcccccggtacacggcggtgtcagcag 1620
QY 1621 tacgttagtcttgacctgcgcgcctggaggtgcggcgggggtgcgcgcagggcctgc 1680
Db 1621 tacgttagtcttgacctgcgcgcctggaggtgcggcgggggtgcgcgcagggcctgc 1680
QY 1681 gcttctggaaccgcttctctcccaaatgtctcagcgcaccc 1722
Db 1681 gcttctggaaccgcttctctcccaaatgtctcagcgcaccc 1722

RESULT 3
AAQ99002
ID AAQ99002 standard; DNA; 2256 BP.
XX AC AAQ99002;
XX AC AAQ99002;
XX 31-MAR-1996 (first entry)
XX Human acetylcholinesterase (AChE) gene.
XX Acetylcholinesterase; acetyl cholinesterase; EC-3.1.1.7;
XX chromosome-7q22; acetylcholine-hydrolyzing enzyme; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX FT 160..2206
XX FT /*tag= a
XX W09523158-Al.
XX 31-AUG-1995.
XX 28-FEB-1995; 95WO-US02806.
XX 09-JAN-1995; 95US-0370156.
XX 28-FEB-1994; 94US-0202755.
XX (KOHN/) KOHN K I.
XX (YISS) YISSUM RES & DEV CO.
XX Shani M, Soreq H, Zakut H;
XX WPI; 1995-311499/40.
XX P-PSDB; AAR80726.
XX Alternative forms of human acetyl cholinesterase (ChE) gene -
XX expressed in transgenic animal assay system for evaluating anti-ChE
XX activity of organo:phosphate(s), etc. or as model of ChE imbalance
XX Claim 3; Fig.1A; 55pp; English.
XX This DNA sequence encoding human acetylcholinesterase is useful
XX for producing transgenic animals which express AChE. The
XX transgenic animals are in turn useful as an assay system for
XX determining the anti-ChE activity of organophosphates, carbamates,
XX anti-ChE drugs, plant glycoalkaloids and snake venoms. This gene
XX contains a promoter region, 6 exons (E1 to E6) and 4 introns (I1-
XX I4).
XX Sequence 2256 BP; 390 A; 757 C; 680 G; 429 T; 0 other;
SQ

Query Match 99.8%; Score 1722; DB 16; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagggcccgagctgtctgtgcacagccttccctggcttcccccactcttctctc 60
Db 160 atgagggcccgagctgtctgtgcacagccttccctggcttcccccactcttctctc 219

Qy 61 ctccctgagctccctgggtggagagtgaggctgagggccggaggatgcagagctgctg 120
Db 220 ctccctggtccctccgggtggaggagtgaggctgagggccggaggatgcagagctgctg 279
Qy 121 gtacgggtgcgtggggccggctgcggggcatttcgctgaagacccccggggccctgtc 180
Db 280 gtacgggtgcgtggggccggctgcggggcatttcgctgaagacccccggggccctgtc 339
Qy 181 tctgctttctcctgggcatccctcttcgggagccaccccatgggacccctgcgtttctgtcca 240
Db 340 tctgctttctcctgggcatccctcttcgggagccaccccatgggacccctgcgtttctgtcca 399
Qy 241 ccgagagcccaagcagcttgatcaggggtgatacagctacaaccttcacagagtctgc 300
Db 400 ccgagagcccaagcagcttgatcaggggtgatacagctacaaccttcacagagtctgc 459
Qy 301 taccaatatgtggacacccctataccaggttttgaggacccagagatgtgaaccccaac 360
Db 460 taccaatatgtggacacccctataccaggttttgaggacccagagatgtgaaccccaac 519
Qy 361 cgtgagctgagcaggagctgctgtacctcaacgtgtggaacaccatacccccgctctaca 420
Db 520 cgtgagctgagcaggagctgctgtacctcaacgtgtggaacaccatacccccgctctaca 579
Qy 421 tccccccacctgctcctgctgagatctatgggggtgagcttctacagtggggcctctcc 480
Db 580 tccccccacctgctcctgctgagatctatgggggtgagcttctacagtggggcctctcc 639
Qy 481 ttggacgtgtacgattggcgtctcttggtacagcccgagagactgtctggttccatg 540
Db 640 ttggacgtgtacgattggcgtctcttggtacagcccgagagactgtctggttccatg 699
Qy 541 aactacccgggtggagacctttggcttctcctggcctgcggggagccgagagccccgggc 600
Db 700 aactacccgggtggagacctttggcttctcctggcctgcggggagccgagagccccgggc 759
Qy 601 aatgtgggtctcctggatcagagctgacctgcaagtgggtgcaggaacacgtggcagcc 660
Db 760 aatgtgggtctcctggatcagagctgacctgcaagtgggtgcaggaacacgtggcagcc 819
Qy 661 ttcgggggtgaccgcgacatcagctgttttggggagagcgcgggagccgctcggtg 720
Db 820 ttcgggggtgaccgcgacatcagctgttttggggagagcgcgggagccgctcggtg 879
Qy 721 ggcatacctctgcttcccgcacagccggggcctgttccacagggccgctgctgcagagc 780
Db 880 ggcatacctctgcttcccgcacagccggggcctgttccacagggccgctgctgcagagc 939
Qy 781 ggtcccccgaatgacctggccacagtgagcattggagagagcccgctcgagggccacg 840
Db 940 ggtcccccgaatgacctggccacagtgagcattggagagagcccgctcgagggccacg 999
Qy 841 cagctggcccacctgtggtgcttctccagcgccgacctgtgggaaatgacacagagctg 900
Db 1000 cagctggcccacctgtggtgcttctccagcgccgacctgtgggaaatgacacagagctg 1059
Qy 901 gtacgctgctctggacacgaccagcgaggtcctgtgtgaacccagaaatggcagctgctg 960
Db 1060 gtacgctgctctggacacgaccagcgaggtcctgtgtgaacccagaaatggcagctgctg 1119
Qy 961 cctcaagaaacgcttctccggttctcctgctgctgtgtagatggagacctctcag 1020
Db 1120 cctcaagaaacgcttctccggttctcctgctgctgtgtagatggagacctctcag 1179
Qy 1021 gacacccagagggcctcatcaacgcgggagaccttccacggcctgacaggtgctggtgggt 1080
Db 1180 gacacccagagggcctcatcaacgcgggagaccttccacggcctgacaggtgctggtgggt 1239
Qy 1081 gtggtgaagagatgagggctgtattttctgtttacggggccccagggcttcagcaaaagac 1140
Db 1240 gtggtgaagagatgagggctgtattttctgtttacggggccccagggcttcagcaaaagac 1299
Qy 1141 aacgagctctcatcagcggcgagcttctgcccgggtgcgggttcccccag 1200

Db 1300 aacgagctctcatcagcggcgagttcctggccgggtgcgggttcggggttccccag 1359
Qy 1201 gtaagtaccctggcagccgagagctgtgctgctgattacacagagctgctgcatacccgag 1260
Db 1360 gtaagtaccctggcagccgagagctgtgctgctgattacacagagctgctgcatacccgag 1419
Qy 1261 gaccgggcacgctgagggagccctgagcagatgtggtgggagacacaaatgtcgtg 1320
Db 1420 gaccgggcacgctgagggagccctgagcagatgtggtgggagacacaaatgtcgtg 1479
Qy 1321 cccatggcccaagctggctggcgagctgctcccaaggtgccgggttctacgctcacgtc 1380
Db 1480 cccgtggcccaagctggctggcgagctgctcccaaggtgccgggttctacgctcacgtc 1539
Qy 1381 ttggaacacctgcttccacgctctcctggccctgtggtgagtggggtgccccacagctac 1440
Db 1540 ttggaacacctgcttccacgctctcctggccctgtggtgagtggggtgccccacagctac 1599
Qy 1441 gagatcgagttcatcttttgggacccccctggacccctctcgaacctcgaacctcagagag 1500
Db 1600 gagatcgagttcatcttttgggacccccctggacccctctcgaacctcagagagag 1659
Qy 1501 aaaatcttcgccagcagctgatactgatactggggcgaacctttgccgacaggggatccc 1560
Db 1660 aaaatcttcgccagcagctgatactgatactggggcgaacctttgccgacaggggatccc 1719
Qy 1561 aatgagccccgagaccccccaagcccccaaatggccccgtacacgcgggggtcagcag 1620
Db 1720 aatgagccccgagaccccccaagcccccaaatggccccgtacacgcgggggtcagcag 1779
Qy 1621 tacgttagctggaacctgcccgcgtggaggtggcggggggtgctgcgccccagggcctgc 1680
Db 1780 tacgttagctggaacctgcccgcgtggaggtggcggggggtgctgcgccccagggcctgc 1839
Qy 1681 gcccttgaacaccttctcccaaatgtctcagcgcacc 1722
Db 1840 gcccttgaacaccttctcccaaatgtctcagcgcacc 1881

RESULT 4
AAS17547
ID AAS17547 standard; DNA; 5767 BP.
XX
AC AAS17547;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM034.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SORE H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
PI Mor T, Soreq H, Arntzen C, Mason H;

|||||
Db 2404 cccgagaccacagggcccccacatggccccgtacacggcgggggtccagcagtcagttt 2463
Qy 1627 agtctggacctgcgcgcgtgaggtgaggtggcgggggtgcgcgccagggctgcgccttc 1686
Db 2464 agtctggacctgcgcgcgtgaggtgaggtggcgggggtgcgcgccagggctgcgccttc 2523
Qy 1687 tggaaacgcttcctcccaaatgtctcagcgc 1718
Db 2524 tggaaacgcttcctcccaaatgtctcagcgc 2555
|||||
RESULT 5
AAS17548
ID AAS17548 standard; DNA; 14446 BP.
XX AC AAS17548;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM036.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM036;
KW circular; ds.
XX
OS Synthetic.
XX
PN W0200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001W0-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASON/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
PI Mor T, SoREQ H, Arntzen C, Mason H;
XX
XX WPI; 2002-055120/07.
XX
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
XX Claim 11; Page 32-41; 42pp; English.
XX
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable

CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM036, the pOPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;

Query Match 99.2%; Score 1712; DB 24; Length 14446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ccccgcaagtgtctgtgcacacgccttcctcggttcccaactctctctctctctc 66
Db 148 ccccgcaagtgtctgtgcacacgccttcctcggttcccaactctctctctctc 207
Qy 67 tggctcctgggtgagagatggggctgagggccggagagatgcagagctctggtgacg 126
Db 208 tggctcctgggtgagagatggggctgagggccggagagatgcagagctctggtgacg 267
Qy 127 gtgcgtggggccgcgcgtgcggggcattcgcctgaagaccccgggggccctgtctctgct 186
Db 268 gtgcgtggggccgcgcgtgcggggcattcgcctgaagaccccgggggccctgtctctgct 327
Qy 187 ttctgtggcatccctcttgcgaggaccccatgggaccccgctcttcttccaccggag 246
Db 328 ttctgtggcatccctcttgcgaggaccccatgggaccccgctcttcttccaccggag 387
Qy 247 cccaagcagccttggtcaggggtggtagacgctcaaaccttccagagtgctgtctataccaa 306
Db 388 cccaagcagccttggtcaggggtggtagacgctcaaaccttccagagtgctgtctataccaa 447
Qy 307 tatgtgacacccctatacccaaggttttgaggccacccagagatgtggaaccccaacgctgag 366
Db 448 tatgtgacacccctatacccaaggttttgaggccacccagagatgtggaaccccaacgctgag 507
Qy 367 ctgagcgaggagactccctgtacctcaacggttggaacacacatacccccgccctacatcccc 426
Db 508 ctgagcgaggagactccctgtacctcaacggttggaacacacatacccccgccctacatcccc 567
Qy 427 acccctgtcctcgtctgagatctatgggggtgctcttctacagtggggctcctctcttgac 486
Db 568 acccctgtcctcgtctgagatctatgggggtgctcttctacagtggggctcctctcttgac 627
Qy 487 gtgtacgatggcgccttcttgggtacagcccgagagagactgtgctggttccatgaactac 546
Db 628 gtgtacgatggcgccttcttgggtacagcccgagagagactgtgctggttccatgaactac 687
Qy 547 cgggtggagccttggcttccctgcccgtgcggggagccgagagcccccgggcaatgtg 606
Db 688 cgggtggagccttggcttccctgcccgtgcggggagccgagagcccccgggcaatgtg 747
Qy 607 ggtcctctgagcagagggctgcccctgcagtggggtgaggaagaacgtgagcctctcggg 666
Db 748 ggtcctctgagcagagggctgcccctgcagtggggtgaggaagaacgtgagcctctcggg 807
Qy 667 ggtgaccgcagatcagtgacgctgtttgggagagcgcgagcgccctcggtgggcatg 726
Db 808 ggtgaccgcagatcagtgacgctgtttgggagagcgcgagcgccctcggtgggcatg 867
Qy 727 cactgtgttccccgcacagccggggcgtgttccacagggcccgtgtcgtcagagcgggtgcc 786
Db 868 cactgtgttccccgcacagccggggcgtgttccacagggcccgtgtcgtcagagcgggtgcc 927
Qy 787 cccaatgacctgggcccacaggtgggcatggagagggcccctcaggggcccacgcagctg 846
Db 928 cccaatgacctgggcccacaggtgggcatggagagggcccctcaggggcccacgcagctg 987
Qy 847 gccacacttgggctgtccctccagggcggaactggtgggaatgacacagagctggtgagcc 906
Db 988 gccacacttgggctgtccctccagggcggaactggtgggaatgacacagagctggtgagcc 1047

```
QY 907 tgccttcgagacgacacgagctcctcgttgaaacacgaatggcaactgctgctcaa 966
DB 1048 tgccttcgagacgacacgagctcctcgttgaaacacgaatggcaactgctgctcaa 1107
QY 967 gaaacgcttcctcgttcctcgtcgtcgttgtagatggagacttctcagtgacac 1026
DB 1108 gaaacgcttcctcgttcctcgtcgtcgttgtagatggagacttctcagtgacac 1167
QY 1027 ccagagccctcatcaacgaggagacttccacggcctcaggtcgtggtggtggtg 1086
DB 1168 ccagagccctcatcaacgaggagacttccacggcctcaggtcgtggtggtggtg 1227
QY 1087 aaggatgaggtcgtattttctgtttacggggcccccaggttcaagaaacacagc 1146
DB 1228 aaggatgaggtcgtattttctgtttacggggcccccaggttcaagaaacacagc 1287
QY 1147 tctctcatcagcgggcagctcctcggccgggtgctgggttccccaggttaagt 1206
DB 1288 tctctcatcagcgggcagctcctcggccgggtgctgggttccccaggttaagt 1347
QY 1207 gacctggcagcgaggtcgttctcgtcattacacagactggctgcattcccgagacc 1266
DB 1348 gacctggcagcgaggtcgttctcgtcattacacagactggctgcattcccgagacc 1407
QY 1267 gacgctcagggagccctcagcagtgatggtggtggcgacacaaatcgtgtgccc 1326
DB 1408 gacgctcagggagccctcagcagtgatggtggtggcgacacaaatcgtgtgccc 1467
QY 1327 gccagctggctggcgagctggctgccaggtgcccggtctacgcctcctttaa 1386
DB 1468 gccagctggctggcgagctggctgccaggtgcccggtctacgcctcctttaa 1527
QY 1387 caccgtgttccacgctcctcggccctcgtggtgggtgggtgggtgggtgggtgggt 1446
DB 1528 caccgtgttccacgctcctcggccctcgtggtgggtgggtgggtgggtgggtgggt 1587
QY 1447 gagttaatttgggataccctcggacccctcgtggtgggtgggtgggtgggtgggtgg 1506
DB 1588 gagttaatttgggataccctcggacccctcgtggtgggtgggtgggtgggtgggtgg 1647
QY 1507 ttccgcccagcagctgactgactgactgactgactgactgactgactgactgactg 1566
DB 1648 ttccgcccagcagctgactgactgactgactgactgactgactgactgactgactg 1707
QY 1567 ccccgagaccccaaggcccaaatggcccccgtaacacgacgggggctcagcagtc 1626
DB 1708 ccccgagaccccaaggcccaaatggcccccgtaacacgacgggggctcagcagtc 1767
QY 1627 agtctggacctcggcgcgtgaggtgaggtgaggtgaggtgaggtgaggtgaggt 1686
DB 1768 agtctggacctcggcgcgtgaggtgaggtgaggtgaggtgaggtgaggtgaggt 1827
QY 1687 tggaaacgcttctcctcccaattgctcagcgc 1718
DB 1828 tggaaacgcttctcctcccaattgctcagcgc 1859
```

RESULT 6

AAQ05998

ID AAQ05998 standard; DNA; 2253 BP.

XX

AC AAQ05998;

XX

DT 16-JAN-1991 (first entry)

XX

DE Sequence encoding human acetylcholinesterase (hAChE).

XX

KW Organophosphorous poisoning; OP; cancer; leukaemia;

KW megakaryocytopoiesis; ovarian cancer; ds.

XX

OS Homo sapiens.

XX

PN EP388906-A.

```
XX 26-SEP-1990.
PD
XX 20-MAR-1990; 90EP-0105274.
PF
XX 21-MAR-1989; 89IL-0089703.
PR
XX (YISS ) YISSUM RES DEV CO.
PA
XX Soreq H, Zakut H;
PI
XX WPI: 1990-291865/39.
DR P-PSDB; AAR06989.
XX
PT Human acetylcholinesterase DNA and prodn. of recombinant hAChE -
PT for treatment of organo-phosphorous poisoning and diagnosis of haemo-
PT cytopoeitic disorders and ovarian carcinomas.
XX
PS Claim 5; Page 25; 47pp; English.
XX
CC Gene product is useful as an active pharmacological component for the
CC prophylaxis and treatment of organophosphorous poisoning, and post-
CC surgical apnea due to succinylcholine administration.
CC cDNA probe to the sequence may be used in diagnosis of various
CC leukemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
XX
SQ Sequence 2253 BP; 390 A; 740 C; 694 G; 429 T; 0 other;
```

Query Match 72.6%; Score 1252; DB 11; Length 2253;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1552; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 165 ccccgggggccctgctctgcttctcctggcgaccccccttgcggagccaccatgggacc 224
DB 321 ccccgggggccctgctctgcttctcctggcgaccccccttgcggagccaccatgggacc 380
QY 225 ccgtgcttctgcacccggagcccaagcagccttggtcagggggtgtagacgtacaac 284
DB 381 ccgtgcttctgcacccggagcccaagcagccttggtcagggggtgtagacgtacaac 440
QY 285 cttccagagctcgtctacaaatgtagacacccatataccacaggttttgggggaccga 344
DB 441 cttccagagctcgtctacaaatgtagacacccatataccacaggttttgggggaccga 500
QY 345 gatgtggaaccccaacccgtgagctgagcagagactcctgtacctcaacgtgtggacac 404
DB 501 gatgtggaaccccaacccgtgagctgagcagagactcctgtacctcaacgtgtggacac 560
QY 405 ataccgccggcctacataccccacccctgtcctgctggtatctatgggggtggtctcta 464
DB 561 ataccgccggcctacataccccacccctgtcctgctggtatctatgggggtggtctcta 620
QY 465 cagtggggctcctcttgagcgtgtacatggcgccttcttggtacagccgagagagac 524
DB 621 cagtggggctcctcctcttgagcgtgtacatggcgccttcttggtacagccgagagagac 580
QY 525 tgtgctggtgtccatgaactacccgggtgggagccttcttgctcctgcgcctgcccgggag 584
DB 681 tgtgctggtgtccatgaactacccgggtgggagccttcttgctcctgcgcctgcccgggag 740
QY 585 ccgagaggccccgggcaatgtggtgtctcctgatacagagctggccctcagtggtgtca 644
DB 741 ccgagaggccccgggcaatgtggtgtctcctgatacagagctggccctcagtggtgtca 800
QY 645 ggagaaactggcagccttcgggggtgacccacatcagtgacgctgtttgggagagagcgc 704
DB 801 ggagaaactggcagccttcgggggtgacccacatcagtgacgctgtttgggagagagcgc 860
QY 705 gggagccgctcgggtgggcatgcacactgcttccccgccagccggggcctgttccacag 764
DB 861 gggagccgctcgggtgggcatgcacactgcttccccgccagccggggcctgttccacag 920
```


Qy	765	ggcgcgtctgcagagcggtgcccccaatggaccttgggccacgdtgggcccatgtggccatgtggagaggc	824
Db	921	ggcgcgtctgcagagcggtgcccccaaTggaccttgggccacgdtgggcccatgtggagaggc	980
Qy	825	ccgtccaggggccacgcagctaggcccacttgtggcgtctcctccaaggcgcactggtagg	884
Db	981	ccgtccaggggccaccacagctggccacacttgtggcgtctcctccaaggcgcactggtagg	1040
Qy	885	gaatgacacagagtggtagctgcttcggacacgaccagcgagtgctctggTgaacca	944
Db	1041	gaatgacacagagtggtagctgcttcggacacgaccagcgagtgctctggTgaacca	1100
Qy	945	cgaattggcacgtgctgcctcaagaagcgtttccogTttcctctctgctcgtctggtaga	1004
Db	1101	cgaattggcacgtgctgcctcaagaagcgtttccogTttcctctctgctcgtctggtaga	1160
Qy	1005	tgaagacttctcagtgacaccccagagccctcaacaacgaggagacitccaagcct	1064
Db	1161	tgaagacttctcagtgacaccccagagccctcaacaacgaggagacitccaagcct	1220
Qy	1065	gcaggTgctTggTgggtgaagatTgagggtcgttatTTctTgTtttacggggcccc	1124
Db	1221	gcaggTgctTggTgggtgaagatTgagggtcgttatTTctTgTtttacggggcccc	1280
Qy	1125	agcttcaacaaagacaaagctctcatcagccggccagTtcbtgcccggggtgcg	1184
Db	1281	agcttcaacaaagacaaagctctcatcagccggccagTtcbtgcccggggtgcg	1340
Qy	1185	ggTcggggTtccccagTtaagTgacctggccagccagagctgtgctcgtcatcacaga	1244
Db	1341	ggTcggggTtccccagTtaagTgacctggccagccagagctgtgctcgtcatcacaga	1400
Qy	1245	ctggctTgcatccccagagaccggccacgctgtagggaggccctgagcgatgtTgTgggcga	1304
Db	1401	ctggctTgcatccccagagaccggccacgctgtagggaggccctgagcgatgtTgTgggcga	1460
Qy	1305	ccacaatgtcgtgtgccccgtggccacgtggctgggcgactggctgccagaggtgcceg	1364
Db	1461	ccacaatgtcgtgtgccccgtggccacgtggctgggcgactggctgccagaggtgcceg	1520
Qy	1365	ggTctacgctaagctctTgaaacacgTgcttccacgctctcctTggccccctgtgatatggg	1424
Db	1521	ggTctacgctaagctctTgaaacacgTgcttccacgctctcctTggccccctgtgatatggg	1580
Qy	1425	ggTgccccacggtacagatacgattcatctTtggaTccccctggacccccctctcgaaa	1484
Db	1581	ggTgccccacggtacagatacgattcatctTtggaTccccctggacccccctctcgaaa	1640
Qy	1485	ctacacggcagaggagaaaaattcttcgcacgcgactgatgcgatactTggccaaactTgc	1544
Db	1641	ctacacggcagaggagaaaaattcttcgcacgcgactgatgcgatactTggccaaactTgc	1700
Qy	1545	ccgcacaggggatTcccaatTgaacccccgagacccaagccccacaaTggcccccgTcac	1604
Db	1701	ccgcacagggggaTcccaaatTgaacccccgagacccaagccccacaaTggcccccgTcac	1760
Qy	1605	ggcgggggctcagcagTacgtTagtTgagactgcggccgctggaggtTgcggcggggct	1664
Db	1761	ggcgggggctcagcagTacgtTagtTgagactgcggccgctggaggtTgcggcggggct	1820
Qy	1665	gdcgcgccaggcctgcgccttctgaaocgcttctccccaaaattgctcagcgccacc	1722
Db	1821	gdcgcgccaggcctgcgccttctgaaocgcttctccccaaaattgctcagcgccacc	1878

7 RESULT

AAS17492

AAS17492
ID AAS17492 standard; DNA; 12113 BP.

XX

AC AAS17492;

XX

DT 26-FEB-2002 (first entry)

XX

DE	Human	acetylcholinesterase, ACHE, gene.
XX		
KW	Human; ds;	ACHE; acetylcholinesterase; polymorphic variant; haplotype; neurological disease; Parkinson's disease; Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.2.
KW		
KW		
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	variation	replace (1950,C)
FT		/*tag= a
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (2237,T)
FT		/*tag= b
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (4044,T)
FT		/*tag= c
FT		/standard_name= "Single nucleotide polymorphism"
FT	CDS	4095..8113
FT		/*tag= d
FT		/product= "ACHE"
FT	exon	4095..5162
FT		/*tag= e
FT		/number= 2
FT	intron	5163..5508
FT		/*tag= f
FT		/number= 2
FT	variation	replace (4130,T)
FT		/*tag= g
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (4195,A)
FT		/*tag= h
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (4277,C)
FT		/*tag= i
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (4932,G)
FT		/*tag= j
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (4967,T)
FT		/*tag= k
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (5136,A)
FT		/*tag= l
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (5151,A)
FT		/*tag= m
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	5509..5993
FT		/*tag= n
FT		/number= 3
FT	intron	5994..6988
FT		/*tag= o
FT		/number= 3
FT	variation	replace (5987,T)
FT		/*tag= p
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (6831,A)
FT		/*tag= q
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	6989..7158
FT		/*tag= r
FT		/number= 4
FT	intron	7159..7991
FT		/*tag= s
FT		/number= 4
FT	exon	7992..8113
FT		/*tag= t
FT		/number= 5
FT	variation	replace (8114,A)
FT		/*tag= u
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace (8200,A)

```
FT      /*tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace (8227,A)
FT      /*tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace (8425,G)
FT      /*tag= x
FT      /standard_name= "Single nucleotide polymorphism"
XX
XX WO200179219-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US11853.
XX
XX 14-APR-2000; 2000US-197173P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX (KAZE/) KAZEMI A.
XX
XX Bentivegna SC, Chew A, Choi JY, Koshy B;
XX
XX WPI: 2002-055248/07.
XX P-PSDB; AAU11231.
XX
XX New polymorphic variants comprising acetylcholinesterase (ACHE)
XX isogene, useful in expressing ACHE protein for use in screening for
XX candidate drugs to treat diseases related to ACHE activity, e.g.
XX neurological diseases or cancer -
XX
XX Claim 21; Fig 1; 79pp; English.
XX
XX The invention relates to a polynucleotide comprising a polymorphic
XX variant of an acetylcholinesterase (ACHE) gene or fragment, protein or
XX complement, the variant comprising an ACHE isogene defined by a haplotype
XX selected from haplotypes 1-20 listed in the specification. Also included
XX are methods for haplotyping and genotyping the ACHE gene of an
XX individual, a method for predicting a haplotype pair for the ACHE gene of
XX an individual, a method for identifying an association between a trait
XX and at least one haplotype or haplotype pair of ACHE gene, recombinant
XX nonhuman organisms transformed or transfected with the polynucleotide
XX where the organism expresses ACHE protein encoded by the first
XX nucleotide sequence or encoded by the polymorphic variant sequence,
XX an isolated antibody specific for and immunoreactive with ACHE,
XX a method of screening for drugs targeting the polypeptide contacting ACHE
XX polymorphic variant with a candidate agent and assaying for binding
XX activity, a computer system for storing and analysing polymorphism data
XX for ACHE gene and a genome anthology for ACHE gene which comprises ACHE
XX isogenes defined by haplotypes 1-20 given in the specification.
XX The Polymorphisms are useful for studying the biological function of
XX ACHE as well as in identifying drugs targeting this protein for the
XX treatment of disorder related to its abnormal expression or function.
XX The polymorphic variants may also be used in screening for compounds
XX targeting ACHE to treat a specific condition or disease predicted to be
XX associated with ACHE activity e.g. neurological diseases (e.g.
XX Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
XX and tumours. The ACHE gene maps to human chromosome 7q22. The
XX present sequence is the ACHE gene.
XX
XX Sequence 12113 BP; 2215 A; 3836 C; 3766 G; 2296 T; 0 other;
```

```
Query Match      62.0%; Score 1070; DB 24; Length 12113;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 atgagggcccccagtgctgtgtgacacgccttcctcctgattcccccactcttcctc 60
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4095 atgagggcccccagtgctgtgtgacacgccttcctcctgattcccccactcttcctc 4154

QY      61 ctcccttggtcctcctggtgagagtggtgggctgagggccgggagagatgacagctgctg 120
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4155 ctcccttggtcctcctggtgagagtggtgggctgagggccgggagagatgacagctgctg 4214
```

```
QY      121 gtgacggtgcgtggggggccggctgcggggcattcgcctcaagaccgccggggccctgtc 180
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4215 gtgacggtgcgtggggggccggctgcggggcattcgcctcaagaccgccggggccctgtc 4274

QY      181 tctgctttctcctgggcatcccttttcggagcaccacccatggagaccctcgtcttcttgcca 240
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4275 tctgctttctcctgggcatcccttttcggagcaccacccatggagaccctcgtcttcttgcca 4334

QY      241 ccggagcccaagcagccttggttcaggggtggtagagcgtacacacctccacagatgctctgc 300
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4335 ccggagcccaagcagccttggttcaggggtggtagagcgtacacacctccacagatgctctgc 4394

QY      301 taccataatgtggacaccctctaccctgacacgtgtggacacacacaccccccgcctacac 360
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4395 taccataatgtggacaccctctaccctgacacgtgtggacacacacaccccccgcctacac 4454

QY      361 cgtgagctgagcagagagactgctgtaccctcaacgtgtggacacacacaccccccgcctacac 420
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4455 cgtgagctgagcagagagactgctgtaccctcaacgtgtggacacacacaccccccgcctacac 4514

QY      421 tccccacccctcctcgtctgtggtatctatgggggtggtctctacagtggggctcctccc 480
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4515 tccccacccctcctcgtctgtggtatctatgggggtggtctctacagtggggctcctccc 4574

QY      481 ttgagcgtgtacagatggccgcttcttgggtacagggccgagagagactgtctgtgtccatg 540
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4575 ttgagcgtgtacagatggccgcttcttgggtacagggccgagagagactgtctgtgtccatg 4634

QY      541 aactacccgggtgggagaccccttggctcctcctgcccgggggagcagagagcccccgggc 600
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4635 aactacccgggtgggagaccccttggctcctcctgcccgggggagcagagagcccccgggc 4694

QY      601 aatgtgggtctcctgcatcagagcgtggccctcagtggtgcagagacgtggcagcc 660
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4695 aatgtgggtctcctgcatcagagcgtggccctcagtggtgcagagacgtggcagcc 4754

QY      661 ttcgggggtgaccccgacatcagtgacgctgttttggggagagcgcgggagcgcctcggtg 720
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4755 ttcgggggtgaccccgacatcagtgacgctgttttggggagagcgcgggagcgcctcggtg 4814

QY      721 ggcatacactgctgtcccgcgccacggccggtcctttccacagggccgtgctgcagagc 780
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4815 ggcatacactgctgtcccgcgccacggccggtcctttccacagggccgtgctgcagagc 4874

QY      781 ggtgcccccaatggacctgggcccacgggtggcatgggagagcccgctgcagggccacg 840
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4875 ggtgcccccaatggacctgggcccacgggtggcatgggagagcccgctgcagggccacg 4934

QY      841 cagctggccacacttggtggtctcctccaggcggcactggtgggaatgacacagagctg 900
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4935 cagctggccacacttggtggtctcctccaggcggcactggtgggaatgacacagagctg 4994

QY      901 gtacgctgcttcggacacacacacgcagcagctcctggtgaacacacgaatggcagctgctg 960
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4995 gtacgctgcttcggacacacacacgcagcagcagctcctggtgaacacacgaatggcagctgctg 5054

QY      961 cctcaagaagcgtcttcgggttctcctcctcctcctcctcctcctcctcctcctcctcag 1020
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      5055 cctcaagaagcgtcttcgggttctcctcctcctcctcctcctcctcctcctcctcag 5114

QY      1021 gacaccccgagggccctcctcctcctcctcctcctcctcctcctcctcctcaggt 1070
DB      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      5115 gacaccccgagggccctcctcctcctcctcctcctcctcctcctcctcctcaggt 5164
```

```
RESULT      8
ID      AAQ05999
XX      AAQ05999 standard; DNA; 1800 BP.
AC      AAQ05999;
XX
XX      16-JAN-1991 (first entry)
DT
```

```

XX DE Sequence encoding foetal human acetylcholinesterase (hAChE).
XX KW Organophosphorous poisoning; OP; cancer; leukaemia;
XX KW megakaryocytopoiesis; ovarian cancer; ds.
XX OS Homo sapiens.
XX FH
XX FT Location/Qualifiers
XX FT 1..1500
XX FT /*tag= a
XX FT 1501..1800
XX FT /*tag= b
XX PN EP388906-A.
XX PD
XX PD 26-SEP-1990.
XX PF 20-MAR-1990; 90EP-0105274.
XX PR 21-MAR-1989; 89IL-0089703.
XX PA (YISS ) YISSUM RES DEV CO.
XX PI Soreq H, Zakut H;
XX PI WPI; 1990-291865/39.
XX DR P-PSDB; AAR06990.
XX DR
XX PT Human acetylcholinesterase DNA and prodn. of recombinant hAChE -
XX PT for treatment of organo-phosphorous poisoning and diagnosis of haemo-
XX PT cytopoeitic disorders and ovarian carcinomas.
XX PS Disclosure; Fig 1c; 47pp; English.
XX CC hAChE is useful as an active pharmacological component for the
XX CC prophylaxis and treatment of organophosphorous poisoning, and
XX CC post-surgical apnea due to succinylcholine administration.
XX CC CDNA probe to the sequence may be used in diagnosis of various
XX CC leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
XX CC Bases given as N are obscured in the patent specification.
XX SQ
XX Sequence 1800 BP; 330 A; 602 C; 539 G; 326 T; 3 other;

Query Match 28.3%; Score 489; DB 11; Length 1800;
Best Local Similarity 99.5%; Pred. NO. 5.1e-217;
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 523 actgtgtgtgttccatgaactaccgggtggagcctttgtgcttctgcccctgcgggg 582
DB 178 actgtgtgtgttccatgaactaccgggtggagcctttgtgcttctgcccctgcgggg 237
QY 583 agccgagagggcccggaatgtggttctcctggatcagagctggccctgcagtgggtg 642
DB 238 agccgagagggcccggaatgtggttctcctggatcagagctggccctgcagtgggtg 297
QY 643 caggagaaactgtgcagccttcctggggtgacccgacatcagtcgtgttggggagagc 702
DB 298 caggagaaactgtgcagccttcctggggtgacccgacatcagtcgtgttggggagagc 357
QY 703 gcggagagcgcctcgttgggcatcacctgtgttcccgcccgagcgggctgttccac 762
DB 358 ccgggagagcgcctcgttgggcatcacctgtgttcccgcccgagcgggctgttccac 417
QY 763 agggccgtgtcgtcagagcgtgtcccccgaatgacccctgggcccacaggtggcatggagag 822
DB 418 agggccgtgtcgtcagagcgtgtcccccgaatgacccctgggcccacaggtggcatggagag 477
QY 823 gccctcgcagggccacgcagctggccacacttgggtgttccctccagcgccactggt 882
DB 478 gccctcgcagggccacgcagctggccacacttgggtgttccctccagcgccactggt 537

```

```

QY 883 gggatgacacagagctggttagccttcggagacacacagcagcaggtcctgtggaac 942
DB 538 gggatgacacagagctggttagccttcggagacacacagcagcaggtcctgtggaac 597
QY 943 caggaatggcagctgctgcctcaagaaagcgtcttcgggttctccttcgtgctgtgga 1002
DB 598 caggaatggcagctgctgcctcaagaaagcgtcttcgggttctccttcgtgctgtgga 657
QY 1003 gatgagacttctcagtgacaccccgaggccctcatcaacgcgggagacttccacgc 1062
DB 658 gatgagacttctcagtgacaccccgaggccctcatcaacgcggagacttccacgc 717
QY 1063 ctgcaggtgctgtggtgtggtgaaggatgagggctctatttctgtttacggggc 1122
DB 718 ctgcaggtgctgtggtgtggtgaaggatgagggctctatttctgtttacggggc 777
QY 1123 ccaggcttcagcaagacacagctctctcatcagcgggccc 1164
DB 778 ccaggcttcagcaagacacagctctctcatcagcgggccc 819

RESULT 9
AAT34118
ID AAT34118 standard; RNA; 35 BP.
XX
AC AAT34118;
XX
DT 26-FEB-1997 (first entry)
XX
DE Acetylcholinesterase primer, 1522+.
XX
KW Antisense oligonucleotide; BCHE; butyrylcholinesterase; BuChE;
KW AS-BCHE; primary bone marrow cell culture; erythropoietin; primer;
KW cellular differentiation; anti-cholinesterase; ChE; exposure;
KW anti-ChE drug therapy; neurodegenerative disease; hematological disease;
KW carbamate insecticide; cocaine analogue; ss.
XX
OS Synthetic.
XX
PN WO962174-A1.
XX
PD 18-JUL-1996.
XX
PF 11-JAN-1996; 96WO-US00322.
XX
PR 09-JAN-1995; 95US-0370204.
XX
PA (KOHN/) KOHN K I.
XX (YISS ) YISSUM RES & DEV CO.
XX Soreq H, Zakut H;
XX WPI; 1996-342296/34.
XX
XX Screening for pre-disposition to anti-cholinesterase exposure - by
XX analysing serum levels of butyrylcholinesterase or screening DNA
XX for the presence of BuChE allele(s)
XX Example 4; Page 63; 73pp; English.
XX
XX The sequences given in AAT34118-19 are primers which were used to
XX amplify the BCHE (butyrylcholinesterase) gene and the AChE (acetyl-
XX cholinesterase) gene. These primers were used to investigate
XX fetal protection from anti-ChE poisons that may be associated with
XX pregnancy. These primers were used to amplify DNA from placental
XX tissue. Different gestational age showed marked different intensities
XX of signal of each enzyme gene, as did the levels of transcription of
XX each of the enzyme genes at a particular stage. Oligonucleotides
XX amplified using these primers may be used to identify patients with
XX a genetic predisposition for adverse reactions to anti-cholinesterase
XX (ChE) exposure from, e.g. anti-ChE drug therapy for neurodegenerative
XX diseases or hematological diseases, or exposure to carbamate
XX insecticides or cocaine or cocaine analogues.

```


PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465566/50.
XX Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
XX diseases -
PS Disclosure; SEQ ID NO 2032; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AAS41685-AAS42192 represent DNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 8418 BP; 1951 A; 2317 C; 1938 G; 2212 T; 0 other;

Query Match 1.6%; Score 28; DB 22; Length 8418;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 atactggggccaaacttggccgcacaggg 1554
Db 1989 atactggggccaaacttggccgcacaggg 2016

RESULT 12
AAS74821
ID AAS74821 standard; cDNA; 816 BP.
XX
AC AAS74821;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10625.

```
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG10634.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 1; SEQ ID NO 10625; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 816 BP; 191 A; 238 C; 199 G; 188 T; 0 other;
```

```
Query Match 1.6%; Score 27; DB 23; Length 816;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1527 atactggccaacttggccgcacagg 1553
|||||
Db 252 atactggccaacttggccgcacagg 278
```

```
RESULT 13
AAS74822
ID AAS74822 standard; cDNA; 1532 BP.
AC AAS74822;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #10626.
```

```
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG10635.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 1; SEQ ID NO 10626; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 1532 BP; 348 A; 423 C; 441 G; 320 T; 0 other;
```

```
Query Match 1.6%; Score 27; DB 23; Length 1532;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1527 atactggccaacttggccgcacagg 1553
|||||
Db 1337 atactggccaacttggccgcacagg 1363
```

```
RESULT 14
AAD09548
ID AAD09548 standard; cDNA; 1538 BP.
AC AAD09548;
XX 10-SEP-2001 (first entry)
XX Human protease protein-11 (PRTS-11) cDNA.
XX Human; protease protein-11; PRTS-11; cytostatic; hypotensive; antiviral;
```

KW gastrointestinal disorder; anorexia; dysphagia; cardiovascular disorder;
KW atherosclerosis; vasculitis; autoimmune disorder; inflammatory disorder;
KW Alzheimer's disease; cell proliferative disorder; dermatitis; cirrhosis;
KW acquired immune deficiency syndrome; AIDS; neurological disorder; asthma;
KW developmental disorder; epithelial disorder; eczema; dementia; nontropic;
KW neurological disorder; reproductive disorder; infertility; teratogenesis;
KW immunosuppressive; drug screening; actinic keratosis; cardiac; epilepsy;
KW anaemia; antitumour; gene therapy; antibacterial; ss.
XX
OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX CDS 443..676

XX FT /*tag= a

XX FT /product= "Human protease protein-11 (PRTS-11)"

XX WO200146443-A2.

XX

XX PD 28-JUN-2001.

XX PF 19-DEC-2000; 2000WO-US4811.

XX PR 23-DEC-1999; 99US-0172055.

XX PR 21-JAN-2000; 2000US-0177334.

XX PR 28-JAN-2000; 2000US-0178884.

XX PR 02-FEB-2000; 2000US-0179903.

XX PA (INCY-) INCYTE GENOMICS INC.

XX

XX PI Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R, Yue H;

XX PI Nguyen DB, Tang YT, Yao MG, Lal P;

XX

XX DR WPI: 2001-418080/44.

XX DR P-PSDB: AAE04884.

XX

XX PT Novel human protease proteins (PRTS) useful for diagnosing, treating,
XX PT preventing gastrointestinal, cardiovascular, autoimmune/inflammatory,
XX PT cell proliferative disorders associated with abnormal expression of
XX PT PRTS -

XX PS Claim 5; Page 128; 129pp; English.

XX The present sequence is human protease protein (PRTS-11) cDNA. Human PRTS
XX and its nucleic acid molecule are useful for the diagnosis, treatment and
XX prevention of disorders associated with increased or decreased expression
XX of PRTS. Examples of such disorders include, gastrointestinal disorder
XX such as anorexia, dysphagia; cardiovascular disorder such as
XX atherosclerosis, vasculitis; autoimmune/inflammatory disorder such as
XX acquired immune deficiency syndrome (AIDS), asthma; cell proliferative
XX disorder such as actinic keratosis, cirrhosis; developmental disorder
XX such as epilepsy, anaemia; epithelial disorder such as allergic contact
XX dermatitis, eczema; neurological disorder such as Alzheimer's disease,
XX dementia and reproductive disorder such as infertility and teratogenesis.
XX PRTS DNA is useful for creating 'knockin' humanised animals (pigs) or
XX transgenic animals (mice or rats) to model human disease. PRTS DNA is
XX also in useful in gene therapy. PRTS and its immunogenic fragments are
XX useful for screening libraries of compounds in several drug screening
XX assays. PRTS is useful for analysing the proteome of a tissue or
XX cell type.

XX SQ Sequence 1538 BP; 344 A; 451 C; 372 G; 371 T; 0 other;

Query Match 1.6%; Score 27; DB 22; Length 1538;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1527 atactggcccaactttgccacagg 1553

Db 481 atactggcccaactttgccacagg 507

RESULT 15

AAS74456

ID AAS74456 standard; cDNA; 496 BP.

XX

AC AAS74456;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #10260.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

XX WO200175067-A2.

XX

XX PD 11-OCT-2001.

XX

XX PF 30-MAR-2001; 2001WO-US08631.

XX

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX

XX PA (HYSE-) HYSEQ INC.

XX

XX PI Drmanac RT, Liu C, Tang YT;

XX

XX WPI: 2001-639362/73.

XX DR P-PSDB: ABG10269.

XX

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity -

XX

XX PS Claim 1; SEQ ID No 10260; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human

XX diagnostic coding sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX

XX SQ Sequence 496 BP; 104 A; 161 C; 137 G; 94 T; 0 other;

Query Match 1.4%; Score 25; DB 23; Length 496;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 369 gacgcagactgcctgtacctcaac 393

Db 27 gacgcagactgcctgtacctcaac 51

Search completed: August 31, 2002, 22:59:25

Job time: 16866 sec

THIS PAGE BLANK (USPTO)

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 19:59:53 ; Search time 6064.54 Seconds
(without alignments)
3839.079 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgagggcccgagtgctt.....aattgtcagccaccctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_etc:*
9: gb_estc:*
10: gb_estc2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1070	62.0	2080	11 BC001541	BC001541 Homo sapi
2	692	40.1	780	10 BI753192	BI753192 603026066
3	603	35.0	752	10 BI917862	BI917862 603183996
4	569	33.0	732	10 BG708331	BG708331 602672268
5	509	29.5	781	10 BG705869	BG705869 602669361
6	505	29.3	931	10 BI488594	BI488594 603021122
c 7	502	29.1	534	9 AI332425	AI332425 qq25a10.x
8	464	26.9	786	10 BG703349	BG703349 602685108
c 9	460	26.7	589	9 AI654042	AI654042 ty61a12.x
c 10	440	25.5	562	10 BE466391	BE466391 bz21b04.x
c 11	438	25.4	822	9 AI180022	AI180022 qd35e12.x
12	428	24.8	738	10 BI667712	BI667712 603293067
c 13	427	24.8	478	9 AI082805	AI082805 ox74d09.x
14	421	24.4	625	10 BI464924	BI464924 603207402
15	417	24.2	702	10 BI915019	BI915019 603177204
16	412	23.9	574	9 AI207928	AI207928 ap14d04.x
c 17	405	23.5	456	9 AI081892	AI081892 ox77b06.x

c 18	381	22.1	555	9	AI769167	AI769167 wg34f12.x
c 19	372	21.6	577	9	AI802781	AI802781 wf18c08.x
c 20	370	21.4	405	9	AW050645	AW050645 wz19f11.x
21	363	21.0	489	9	AI091215	AI091215 oo22c11.x
c 22	363	21.0	507	9	AW196326	AW196326 xm31h05.x
c 23	352	20.4	990	10	BI489087	BI489087 603021122
c 24	332	19.2	689	10	BG395920	BG395920 602458524
c 25	331	19.2	483	9	AI761619	AI761619 wg66f02.x
c 26	324	18.8	379	9	AI081808	AI081808 ox77e10.x
c 27	321	18.6	474	9	AI761618	AI761618 wg66f01.x
c 28	300	17.4	368	9	AW182170	AW182170 xj71d10.x
c 29	299	17.3	728	10	BG707817	BG707817 602671169
c 30	268	15.5	483	9	AI761609	AI761609 wg66e02.x
31	249	14.4	462	10	W92595	W92595 ze05a05.r1
32	224	13.0	238	10	BM023426	BM023426 ie80c06.y
33	219	12.7	348	10	HI9772	HI9772 yn55f10.r1
34	219	12.7	378	10	H21027	H21027 yn66g03.r1
c 35	218	12.6	330	9	AW337310	AW337310 xw83e11.x
36	216	12.5	514	9	AA032228	AA032228 zf01f07.s
37	215	12.5	396	10	H21132	H21132 yn65b07.r1
38	213	12.3	586	9	AA446656	AA446656 zw89g10.s
39	212	12.3	547	10	BG707892	BG707892 602671083
40	208	12.1	380	10	HI9954	HI9954 yn53c11.r1
41	198	11.5	302	9	AA346027	AA346027 EST52141
c 42	189	11.0	365	10	H21093	H21093 yn65b07.s1
c 43	188	10.9	467	10	N98267	N98267 za28g11.r1
c 44	183	10.6	183	10	BM023149	BM023149 ie80c06.x
c 45	181	10.5	387	9	AA805231	AA805231 of49g04.s

ALIGNMENTS

RESULT 1

LOCUS BC001541 2080 bp mRNA linear HTC 31-JAN-2002
DEFINITION Homo sapiens, clone IMAGE:3453362, mRNA.
ACCESSION BC001541
VERSION BC001541.1 GI:14705895
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2080)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcdpaxil.stanford.edu

R. M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

Dickson.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: INRAK Plate: 4 Row: e Column: 19

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 7710156

This clone has the following problem: no polyA-tail.

Location/Qualifiers

1. .2080

/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:3453362"
/tissue_type="Placenta, chorioncarcinoma"
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 319 a 692 c 651 g 418 t
ORIGIN

Query Match 62.0%; Score 1070; DB 11; Length 2080;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 atgagggccccagtgctgtgtgacacgcttccctggtctcccaactctcttcctc 60
Db 98 ATGAGGCCCCGAGGTGCTGTGTGACACGCTTCCCTGGCTTCCCACTCTCTTCCTC 157
Qy 61 ctctctggtcctgggtgagagtgagggtgagggccgggagagatgcagagctgtg 120
Db 158 CTCCTCTGGCTCTCTGGGTGGAGAGTGGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 217
Qy 121 gtgagggtcgtggggccgctgcgggcatcgcctgaagacccccggggccctgtc 180
Db 218 GTGACGGTCGTGGGGCCGCTGCGGGCATTCGGCTGAAGACCCCCGGGGCCCTGTG 277
Qy 181 tctgtttctgggcatcccccttgcggagccaccatgggaccccgctgctttctgcc 240
Db 278 TCTGCTTCTGGCATCCCTTTGCGGAGCCACCCATGGACCCCGTTCGTTCGCCA 337
Qy 241 ccggagcccaagcagccttggtaagggtggtagacgtacaacctcagagtgctgc 300
Db 338 CCGGAGCCCAAGCAGCTTGGTCAGGGGTGGTAGAGCTACACCTTCAGAGTGTCTGC 397
Qy 301 taccatgtgagacccctataccagatttgagggcaccgagatgtgaacccccaac 360
Db 398 TACCAATATGTGGACACCTATACCCAGGTTTTGAGGGCACCAGATGTGAACCCCAAC 457
Qy 361 cgtgagctgagcagagactcgtgtacctcaacgtgtgacaccataccccgggctaca 420
Db 458 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGACACCATACATACCCGGCTACA 517
Qy 421 tccccacccctgtcctgtgtgattataggggtgggtttctataagtggggctctcc 480
Db 518 TCCCCACCCCTGTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTC 577
Qy 481 ttggagctgtacgatggcgcctcttgggtacaggccgagagactgtggtgtccatg 540
Db 578 TTGGAGCTGTACGATGGCCGCTCTTGTGTACAGGCCGACAGAGACTGTGTGTCCATG 637
Qy 541 aactaccgggtggagccttggcttctcctgcccctgcccgggagcagagagccccgggc 600
Db 638 AACTACGGGTGGAGCCCTTGGCTTCCTGGCCCTGTCGGGGAGCGAGAGCCCGGGC 697
Qy 601 aatgtgggtctctgtgatacagaggtggccctgcagtgaggtgagagaaagtggcagcc 660
Db 698 AATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGCATGTGGGTGCAGAGAACGTGGCAGCC 757
Qy 661 ttccgggggtgaccgacatcagtgacgctgtttggggagagcgcgggagccgctcggtg 720
Db 758 TTCCGGGGGTGACCCGACATCAGTGACCTCTGTTTGGGAGAGCGCGGAGCCGCTCGGTG 817
Qy 721 ggcatacctgtctctcccgccagcggggcctgttccacagggccgtgtgtgcagagc 780
Db 818 GGCATGCACTGTCTGTCCCGCCAGCCGGGGCTGTTCACAGGGCCGTGTGTGCAGAGC 877
Qy 781 ggtgcccccaatggacccttggccacaggttggcattgggagagggcccgtagcagggccacg 840
Db 878 GTGTGCCCCCAATGGACCTTGGGCGCAGGTGGGATGAGGAGGCCCGTGCAGAGGCCAGC 937
Qy 841 cagctggccccacttgttgggtctctctcagcggcgacactggtgggaatgacacagagctg 900
Db 938 CAGCTGGCCCCACCTTGTGGGCTGTCTCCAGCGGGCACTTGTGGGGAATCACACAGAGCTG 997
```

```
Qy 901 gtagcctgcttcggacacgacgagcgaggtcctgtgtgaaccacgaatggcagctgctg 960
Db 998 GTAGCCTGCTTCGGACACGACGACGAGGCTCTGTTGAACCCACCAATGGCAGCTGCTG 1057
Qy 961 cctcaagaagaagctctccgggttctcctctgctgctgtggtagatgagagacttcctcagt 1020
Db 1058 CCTCAAGAAGCGCTCTCCGGTTCTCTCTGTCGCCCTGGTGTAGATGGAGACTTCCCTCAGT 1117
Qy 1021 gacaccccgagggccctcatcaacgcgggagacttcacggcctgcaggt 1070
Db 1118 GACACCCGAGGCGCTCATCAACGGGGAGACTTCCACGGCGCTGCAGGT 1167
```

```
RESULT 2
BI753192
LOCUS 780 bp mRNA linear EST 25-SEP-2001
DEFINITION 603026066F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196256 5',
mRNA sequence.
ACCESSION BI753192
VERSION BI753192.1 GI:15744770
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library prepared by: The I.M.A.G.E. Consortium (LLNL)
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 row: d column: 17
High quality sequence stop: 780.
Location/Qualifiers
1..780
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196256"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI; 6
Site_2: EcoRV (destroyed); RNA source anonymous pool of
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 111 a 255 c 260 g 154 t
ORIGIN
```

```
FEATURES
source
1..780
Query Match 40.1%; Score 692; DB 10; Length 780;
Best Local Similarity 100.0%; Pred. NO. 6.7e-314;
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 atgagggccccagtgctgtgtgacacgcttccctggtctcccaactctcttcctc 60
Db 48 ATGAGGCCCCGAGGTGCTGTGTGACACGCTTCCCTGGCTTCCCACTCTCTTCCTC 107
Qy 61 ctctctggtcctgggtgagagtgagggtgagggccgggagagatgcagagctgctg 120
Db 108 CTCCTCTGGCTCTCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 167
```

Qy	121	gtacggctgagtgaggggccggctgctggggcattcgctgaaagacccccggggccctgtc	180
Db	168	gtacggctggctggggggccggctgctggggcattcgctgaaagacccccggggccctgtc	227
Qy	181	tctgctttcttggaatcccccttgcggagacacccatggagcccgctcgctttctgcca	240
Db	228	tcctgctttcttggaatcccccttgcggagacacccatggagcccgctcgctttctgcca	287
Qy	241	ccggagcccaagcagccttggtgcaggggtggtagacgctacaaccttcagagtgctcgc	300
Db	288	ccggagcccaagcagccttggtgcaggggtggtagacgctacaaccttcagagtgctcgc	347
Qy	301	taccaatatgtgacacccctataccaccaggttttgagggcacgcagatgtggaacccccac	360
Db	348	taccaatatgtgacacccctataccaccaggttttgagggcacgcagatgtggaacccccac	407
Qy	361	cgtgagctgagcagagactcctctgtaacctcaacgctggaacaccacaccccgacctaca	420
Db	408	cgtgagctgagcagagactcctctgtaacctcaacgctggaacaccacaccccgacctaca	467
Qy	421	tccccacccccctgtcctcgtctggatctatgggggttggtctctacagtggggcctctcc	480
Db	468	tccccacccccctgtcctcgtctggatctatgggggttggtctctacagtggggcctctcc	527
Qy	481	ttggacgtgtacatggccgctctcttggtaagccgcagagagactgtgctggttcacatg	540
Db	528	ttggacgtgtacatggccgctctcttggtaagccgcagagagactgtgctggttcacatg	587
Qy	541	aactcacggatggagaccttgcttgccttcgcccctccggggagcagagaccccgagc	600
Db	588	aactcacggatggagaccttgcttgccttcgcccctccggggagcagagaccccgagc	647
Qy	601	aatgtgggtctccttggaatcagaggtcggcctcagtggtgcaggagaaacgtggcagcc	660
Db	648	aatgtgggtctccttggaatcagaggtcggcctcagtggtgcaggagaaacgtggcagcc	707
Qy	661	ttcggggggtgacccgacatcagtgacgctgtt	692
Db	708	ttcggggggtgacccgacatcagtgacgctgtt	739

RESULT	3	
BI917862		
LOCUS	752 bp	linear
DEFINITION	603183996F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248006 5', mRNA sequence.	EST 17-OCT-2001
ACCESSION	BI917862	
VERSION	BI917862.1	GI:16199790
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 752)	
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.	
	Email: ccapbs@email.nih.gov	
	Tissue Procurement: Life Technologies, Inc.	
	cDNA Library Preparation: Life Technologies, Inc.	
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
	DNA Sequencing by: Incyte Genomics, Inc.	
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
	Plate: LLNMI1625	row: p column: 23
	High quality sequence stop: 752.	
FEATURES	Location/Qualifiers	
	1..752	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5248006"	

FEATURES

```

/clone_lib="NIH_MGC_121"
/lab_host="DHI0B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning)". Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
106 a 248 c 251 g 147 t

```

BASE COUNT	106 a	248 c	251 g	147 t
ORIGIN				

Query Match	35.0%; Score 603; DB 10; Length 752;
Best Local Similarity	100.0%; Pred. No. 4.5e-272;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

Qy	1	atgaggccccgcagtgctcgtcacacgacctccctggcttcccactctcttcctc 60
Dd	80	ATGAGGCCCGCGCAGTGCTGCTGCACACGGCTTCCCTGGCTCCCCACTCCTTCCTC 139
Qy	61	ctcctctggctcctcgggtgaggagtgggggctgagggcgcgaggagtgcagagctgct 120
Dd	140	CTCCTCTGGCTCCCTGGGTGGAGAGTAGTGGGGCTGAGGGCCGGAGAGTGCAGAGCTGCTG 199
Qy	121	gtacagtgctggggccggctacggggcattcgctgaagaccgccgggggcccctgct 180
Dd	200	GTGACGGTGTGGGGGGCGGCTCGGGGGCAATTGCCTGTAAAGACCCC GGGGGGCCCTGTC 259
Qy	181	tctgtttcctgggcaatcccccttcgaggagcaccctatgggaccccgtcgctttctgcga 240
Dd	260	TCTGCTTTCCCTGGGCATCCCCCTTTGGSGAGCACCCCATGGGACC CGTCTTCTTGCCA 319
Qy	241	ccggagcccaagcgccttggctcagggggtgtagacgtacaaccttcagagtgctgc 300
Dd	320	CCGGAGGCCAAGCAGCCCTGTGTCTCAGGGGTGTAGACGCTACAACCTTTCAGAGTGTCTGTC 379
Qy	301	taccaatatgtggacaccctatacccaggttttgagggcaccgagatgtggaacccaac 360
Dd	380	TACCAATATGTGGACACCTATACCAGGTTTTGANGGCCACCGAGATGTGAACCCCAAC 439
Qy	361	cgtgagctgagcgaggactgcgttacotcaacgtgtgacaccatacccccggcctaca 420
Dd	440	C GTGAGCTGAGCGAGGAGTGCCTGTACCTAACGTGTGGACACCATACCCCGGCCTACA 499
Qy	421	tccccaccctctcctgcttgatctatgggggtgagcttctacagtggggcctctcc 480
Dd	500	TCCCCACCCCTGTCCTCGCTGTGATCATGGGGTGGCTTCTACAGTGGGGCCTCTCC 559
Qy	481	ttgacgtgtacnattggccgctctcttgtaacagccgagagactgtgctggtgccatg 540
Dd	560	TTGGACGTGTACGATGGCCGCTTCTTGCTACAGGCCGAGAGACTGTGCTGGTGTCCATG 619
Qy	541	aactaccgggtggagacotttggtctcctgcctgcgggagccgagagagccccgggc 600
Dd	620	AAC TACCGGGTGGGAGCCCTTTGGCTTCTGTCCCTGCCGGGAGCCGAGAGCCCCGGGC 679
Qy	601	aat 603
Dd	680	AAT 682

RESULT	4
LOCUS	BG708331
DEFINITION	602672268f1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4795048 5', mRNA sequence.
ACCESSION	BG708331
VERSION	BG708331.1
KEYWORDS	EST.

	732 bp	mRNA	linear	EST 07-MAY-2001
				IMAGE:4795048 5',

```
SOURCE      human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      1. (bases 1 to 732)
TITLE        NIH-MGC http://mgi.nci.nih.gov/
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10677 row: 9 column: 17
High quality sequence stop: 732.
Location/Qualifiers
1. .732
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',
size selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT  101 a 250 c 239 g 142 t
ORIGIN
Query Match      33.0%; Score 569; DB 10; Length 732;
Best Local Similarity 99.8%; Pred. No. 4.1e-256;
Matches 619; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 atgagccccccgagtgctgtcgtcacacgccttccctggtctcccaactctctctc 60
DB 113 ATGAGGCCCCCGCAGTGCTGTCTGCTGCACACGCTTCCCTGCTTCCCACCTCTCTCCTC 172
QY 61 ctctctggtctctggtgagagtgagtgagggctgagggccgggagagtgagagtgctg 120
DB 173 CTCCTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 232
QY 121 gtacagtgctgagggcgagctgagggcattcgctgagaccccgagaccccgagggccctgtc 180
DB 233 GTACAGGTGCTGGGGCGCGCTGCGGGGCTATTCGCTGAAGACACCCCGGGGGCCCTGTC 292
QY 181 tctgcttcttgggacccctcttgggagccacccatggagaccccgagggccctgtc 240
DB 293 TCTGCTTTCTGTGGCATCCCTTTGCGGAGCCACCCATGGGACCCGCTGCTTCTGCTCA 352
QY 241 ccggagcccaagcagcttggtgaggggtggtgagcgcgtacacaccttccagagtgctgc 300
DB 353 CCGGAGCCCAAGCAGGCTTGGTTCAGGGGTGTAGAGGCTACAACTTCCAGAGTGCTGTC 412
QY 301 taccataatgtgacacccctataccaggttttgaggcaccgagagtggaaccccaac 360
DB 413 TACCATAATGTGACACCCCTATACCCAGGTTTGTAGGGCCACCCGAGATGTGGAACCCCAAC 472
QY 361 cgtgagctgagcagagactgctgtacctcaacgtgtggacacacatacccccggcctaca 420
DB 473 CGTGAGCTGAGCAGGACTGCCCTGTACCTCAACGTTGTGGACACCATACATACCCCGGCTACA 532
```

```
QY 421 tccccacccccctgtctcgtctggatctatg9999tggcttctacagtggggcctctccc 480
DB 533 TCCCCACGCCCTGCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCCTCC 592
QY 481 ttgacagtgtacatgagccgctcttctgtacagccgagagagactgtctggtgctcatg 540
DB 593 TTGACAGTGTACGATGCGCGCTTCTTGTACAGCCCGAGAGGACTGTGCTGGTGTCCATG 652
QY 541 aactaccgggtgggagccttttggcttctgcccctgcccgggagccgagagggcccgggc 600
DB 653 AACTACCGGTGGGAGCCTTGGCTTCTGCGCTGCGGGGAGCCGAGAGGCCCGCCGGC 712
QY 601 aatgtgggtctctctggatca 620
DB 713 AATGTGGGTCTCTCTGGATCA 732

RESULT 5
BG705869
LOCUS      602669361F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792046 5',
DEFINITION mRNA sequence.
ACCESSION  BG705869
VERSION     BG705869.1 GI:13980645
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 781)
NIH-MGC http://mgi.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10669 row: j column: 15
High quality sequence stop: 781.
Location/Qualifiers
1. .781
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',
size selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT  110 a 260 c 262 g 149 t
ORIGIN
Query Match      29.5%; Score 509; DB 10; Length 781;
Best Local Similarity 99.8%; Pred. No. 6.1e-228;
Matches 629; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 atagagccccccgagtgctgtgtgcacacgccttccctggtctcccaactctctcctc 60
DB 113 ATGAGGCCCCCGCAGTGCTGTCTGCTGCACACGCTTCCCTGCTTCCCACCTCTCTCCTC 172
```

```
QY 61 ctctctgtgctctggtgagagagtggtggggtcagagggccgggagagatgcagagctgctg 120
|||||
Db 173 CTCTCTGTGGCTCTGTGGTGTGAGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 232
|||||
QY 121 gtacagtgctgtggggccggctgcggggcattcgcctgaagacccccgggggcccctgtc 180
|||||
Db 233 GTACAGGTGGTGGGGCCGGCTGCGGGGCAATTCGCTGAAGACCCCGGGGGCCCTGTC 292
|||||
QY 181 tctgtctctgtggatccctcttggagagccacccatgggagccctcgtcttcttccca 240
|||||
Db 293 TCTGCTTTCTTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCGCTGCTTCTGCCA 352
|||||
QY 241 ccggagcccaagcagccttgtaggggtgtagagcctacaacctccagagtgctgtgc 300
|||||
Db 353 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 412
|||||
QY 301 taccaatatgtgacacccctataccaggttttgagggccagagatgtgaaaccccaac 360
|||||
Db 413 TACCATATGTGACACCCCTATACCCAGGTTTGTGAGGGCACCCAGAGTGTGAACCCCAAC 472
|||||
QY 361 cgtgagctgagcagagactgcctgtacctcaacgtgtggacacacataccccccggcctaca 420
|||||
Db 473 CGTGAGCTGAGCGAGGACTGCCGTGTACCTCAACGTGTGGACACCATACCCCGGCTTACA 532
|||||
QY 421 tccccacccctgtctgtggtatctatgggggtggtctctacagtggggcctctcc 480
|||||
Db 533 TCCTCCCA-CCCTGCTCTGCTGTGATATATGGGGTGGCTTCTACAGTGGGGCTCTCTCC 591
|||||
QY 481 ttgagctgtacagatggcctcttctgtacagccagagagactgtgtgttccatg 540
|||||
Db 592 TTGGACCTGTACGATGACGATGCTTCTGATACAGCCCGAGAGGACTGTGCTGGTGTCCATG 651
|||||
QY 541 aactacgggtggagagccttggcttctcctggccctgcggggagcagagagcccgggc 600
|||||
Db 652 AACTACCGGTGGAGAGCTTGGCTTCTGCTGCTGCTGCGGGAGCCGAGAGCCCGCGGC 711
|||||
QY 601 aatgtgggtcctgtgataagagctggcc 630
|||||
Db 712 AATGTGGTCTCTCTGGATCAGAGGCTGCC 741
|||||
```

```
RESULT 6
BI488594 931 bp mRNA linear EST 28-AUG-2001
LOCUS 60302112F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 5',
DEFINITION mRNA sequence.
ACCESSION BI488594
VERSION BI488594.1 GI:15327822
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 931)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11479 Row: 1 Column: 06
High quality sequence stop: 771.
Location/Qualifiers
1. .931
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

FEATURES

source

RESULT 7

AI332425/c

LOCUS

DEFINITION

AI332425

ACCSSION

VERSION

```
AI332425 534 bp mRNA linear EST 13-FEB-1999
LOCUS q25a10.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1939530
DEFINITION 3', similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);,
mRNA sequence.
ACCSSION AI332425
VERSION AI332425.1 GI:4068984
```

```
/clone="IMAGE:5191829"
Best Local Similarity 29.3%; Score 505; DB 10; Length 931;
Matches 625; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Query Match
Best Local Similarity 99.8%; Pred. No. 4.8e-226;
Matches 625; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 130 cgtggggcggtgcggggcattcgcctgaagacccccggggccctgtctctgtcttc 189
|||||
Db 1 CGTGGGGCGGGCTGCGGGGCAATTCGCTGAAGACCCCGGGGCGCTGTCTGTCTTTC 60
|||||
QY 190 ctggggatcccccttgcggagccacccatgggacccgcgtcgtcttctgcacccggagccc 249
|||||
Db 61 CTGGGCATCCCTTTGGGGAGCCACCCATGGGACCCCGCTGCTTCTGCCACCGAGGCC 120
|||||
QY 250 aagcagccttggtcaggggtgtagagcctacaacctcccagagtgtctgctaccaat 309
|||||
Db 121 AAGCAGCTTGGTCAGGGGTGGTACAGCTACAACCTTCCAGAGTGTCTGTACCAATAT 180
|||||
QY 310 gtggacacccctataccacaggttttgaggggcaccgagatgtggaaccccaacgtgagctg 369
|||||
Db 181 GTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAACCGTGAGCTG 240
|||||
QY 370 agcagagactcctgtacctcaacgtgtggacacacataccccccggcctacatccccccacc 429
|||||
Db 241 AGCGAGGACTGCCCTGTACCTCAAGCTGTGGACACCATACACCCCGGCTTACATCCCGCCACC 300
|||||
QY 430 cctgtcctcgtctgtgattctatgggggtggtctctacagtgggggtcctcctcttg-gacgt 488
|||||
Db 301 CCTGTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCTGTGTGAGCT 360
|||||
QY 489 gtacagtgccgcctcttctgtgtacagggccgagagactgtgtgtgtccatgaactaccg 548
|||||
Db 361 GTACGATGGCGCTTCTTGTGTACAGGGCGAGAGGACTGTGTGTGTCCATGAACCTACCG 420
|||||
QY 549 ggtggagccttctgtcctcctgcgcctgcggggagcggagagagcccggaatgtggg 608
|||||
Db 421 GGTGGAGCCCTTTGGCTTCTCTGCGCCCTGCGGGGAGCGGAGGCCCGGGCAATGTGGG 480
|||||
QY 609 tctcctgagacagagctggccctgcagtgagtgagtgagagagacgtggcagccttcggggg 668
|||||
Db 481 TCTCTGTGATCAGAGGCTGGCCCTGTCAGTGGGTGCAGGAGAACGTGGGAGCCCTTCGGGGG 540
|||||
QY 669 tgacccacacatcagtcagcctgtttggggagagcgcgggagcccgctcgtgtgggcatgca 728
|||||
Db 541 TGACCCGACATCAGTACGCTGTGTTGGGAGAGCGCGGGAGCCCGCTCGGTGGGCATGCA 600
|||||
QY 729 cctgtgtcccccccgccagccggggcc 754
|||||
Db 601 CCTGTCTCCCCCGCCAGCCGGGGCC 626
|||||
```

BASE COUNT

152 a 294 c 306 g 179 t

ORIGIN

```
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 534)
JOURNAL NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 618 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 460.
FEATURES
source
1..534
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1933530"
/clone_lib="Soares_NHMPUL_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, and ss DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 95 a 170 c 173 g 95 t 1 others
ORIGIN
Query Match 29.1%; Score 502; DB 9; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.1e-224;
Matches 502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1119 gccccaggcttcagcaagacaagctctctcatcagccggcgagttcttgcgcgg 1178
Db 502 GGGCCAGGCTTCAGCAAGACAAGAGTCTCTCATCAGCCGGCGGAGTTCCTGGCCGG 443
Qy 1179 ggtcggttcgggttccccaggtaagtacctggcagccggaggtgtgttctgcatta 1238
Db 442 GGTGGGGTTCGGGGTTCGCCAGGTAAAGTACCTGGCAGCCGAGGCTGTGCTCTCAVTA 383
Qy 1239 cacagactggctgctatccggagaccggcagcctgtagggagccctgagcgatgtggt 1298
Db 382 CACAGACTGGCTGCTATCCGAGGACCCGGCAGCTGAGGAGGCCCTGAGCGATGTGCT 323
Qy 1299 gggcagaccacatgctgtgcccgtggccagctggtggcgactggctgcccaggg 1358
Db 322 GGGCCACCAATGTCGTGTGCCCGCTGGCCAGCTGGCTGGCGACNCTGGCTGCCAGGG 263
Qy 1359 tgcgggtgtctacgctacgttctttgaacacgctgttccacgctctctgcccctgtg 1418
Db 262 TGCCCGGGTCTACGGCTACGTCTTTGAACACCGTCTTCCACGCTCTCTGTCGCCCTGTG 203
Qy 1419 gatgggtgtcccgctacgaatcaggttcattctttggatcccccttgagaccctc 1478
Db 202 GATGGGGTGTCCCGCTACGATCAGTTCATCTTTGGGATCCCGCTGAGACCCCTC 143
Qy 1479 tcgaactacacggcagaggagaaatcttcgccagcgactgatcgatactggggcaca 1538
Db 142 TCGAAACTACAGCGCAGAGGAGAAATCTTCGCCAGCGACTGATGCCATCTGGGCCAA 83
```

```
Qy 1539 ctttgcgcgcacagggggtatcccaatgagcccgagccccagcccccaatggcccc 1598
Db 82 CTTTGCCTCCGACAGGGGATCCCAATGAGCCCGAGACCCCAATGAGCCCAATGAGCC 13
Qy 1599 gtacacggcggggggtcagcag 1620
Db 22 GTACACGGCGGGGGCTCAGCAG 1
RESULT 8
LOCUS BG703349 786 bp mRNA linear EST 07-MAY-2001
DEFINITION 602685108F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4817881 5',
mRNA sequence.
ACCESSION BG703349
VERSION BG703349.1 GI:13975591
KEYWORDS EST..
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 786)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM10718 row: o column: 02
High quality sequence stop: 731.
FEATURES
source
1..786
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4817881"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-xhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 110 a 263 c 261 g 152 t
ORIGIN
Query Match 26.9%; Score 464; DB 10; Length 786;
Best Local Similarity 99.6%; Pred. No. 8.2e-207;
Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 atgagggccccgcagtgctgtgcacgcgcttccctggcttcccaactcttctctc 60
Db 113 ATGAGGCCCGCAGCTGTCTGTGTGCACAGCCTTCCCTGGCTTCCCACTCTCTCTC 172
Qy 61 ctctcttggtctctgggtgagggagtgagggtgctgagggccggagagtgagtgctg 120
Db 173 CTCTCTGTGCTCTGCTGGTGGAGAGTGGGGGCTGAGGGCCGGAGATGCAGAGTGTG 232
Qy 121 gtgacggtacgtggggccgctcggggcttcctcctgaagaccggggggcctgtc 180
Db 233 GTGACGGTGTGGGGGGGGGGCTGCGGGGCAFTTCGCTGAAGACCCCGGGGGCCCTGTC
```

QY 181 tetgtttctctggggaatcccccttttgcggagcaaccatcctggagcccgctgcttttctgca 240
|||||
Db 293 TGTGTTCTTCCCTGGGATCCCTTTTCGGAGCCACCACCATGGGACCCGCTTCTTGCCA 352
|||||
QY 241 ccggagcccaagcagcctgtctcagggtgtagagctctacacactccagagtcttc 300
|||||
Db 353 CCGGAGCCCAAGCAGCCCTTGTGTCAGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 412
|||||
QY 301 taccaatatgtggacacccctataccacaggttttggggccacgagatgtggaaccccaac 360
Db 413 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 472
|||||
QY 361 cgtgagctgagcaggagactgctgtactcctcaacgctgtggaacacacatacccccgcgctaca 420
Db 473 CGTGAGCTGAGCGAGGAGCTGCTGTGTACCTCAACGTTGTGGACACCATACCCCGGGCTACA 532
|||||
QY 421 tccccacacccctgtctcgtctgggtatctatgggggtggctctcagtggggctccctcc 480
Db 533 TCCCCATCCCTGTCTCTGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCCTCCCTCC 592
|||||
QY 481 ttggagctgtacgtagggcgctcttctgttacagggcgagagactgtgctggtgccatg 540
Db 593 TTGGACGTGTACGATGGCGCTTCTTGTGTACAGGCGGAGAGACTGTGCTGTGTCCATG 652
|||||
QY 541 aactaccgggtgggagcctttggctt 566
Db 653 AACTACCGGGTGGGAGCCTTTGGCTT 678
|||||

RESULT 9
AI654042/c 589 bp mRNA linear EST 07-MAR-2000
LOCUS ty61a12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2283550 3'
DEFINITION similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);. mRNA
sequence.
ACCESSION AI654042
VERSION AI654042.1 GI:4738021
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 589)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 850 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 461.
Location/Qualifiers
1. .589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2283550"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-132311, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo.

BASE COUNT 106 a 193 c 186 g 101 t 3 others
ORIGIN

Query Match 26.7%; Score 460; DB 9; Length 589;
Best Local Similarity 99.6%; Pred No. 5 9e-205;
Matches 560; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1059 oggcctgcaggtgctgggtggtggaagatgaggcctgattttctggtttacgg 1118
Db 562 CGGCCTCGAGGTGCTGGTGGGTGTGTAAGGATGAGGCTGCTATTTCCTGGTTACGG 503
|||||
QY 1119 ggcacacggcttcacaaagacacagagctctctatcagccgggcccagttcctggccgg 1178
Db 502 GGCCTCAGGCTTTCAGAAAGACACAGAGTCTCTCATCAGCCGGGGCGAGTTCTTGCCCGG 443
|||||
QY 1179 ggtgcgggttcggggttccccaggttaagtgaacctggcagccgaggtgtgtagtctcgatt 1238
Db 442 GGTGGGGTTCGGGGTTCCTCCAGGTAAAGTACCTGGCAGCCGAGGCTGTGGTCTCGATT 383
|||||
QY 1239 cacagactggtgcataccccgagggacccggcacgctgagggagggccctgagcgatgtgt 1298
Db 382 CACAGACTGGCTGCATCCCGAGGACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGT 323
|||||
QY 1299 ggggacacacaaatgctggtggtgcccagtggtggccagctggtgggagactggtgccaggg 1358
Db 322 GGGCCACCAACAATGTGCTGTGTCNCCTGGCCAGCTGGCTGGGGCTGCTGGCTGCCAGGG 263
|||||
QY 1359 tgcccggttctacgctacgtctttgaacacgctgtcttcacgctctctctgccccctgtg 1418
Db 262 TGCCCGGGTCTACGGCTAGTCTTTTGAACCCGCTGCTTCACGCTCTCTCTGGCCCCCTGTG 203
|||||
QY 1419 gatggggtgccccacggctacgagatgagttcatcttttgggataccccctggaccctc 1478
Db 202 GATGGGGTGCCTCCACGGCTACGAGATCAGTTCATCTTTGGGATCCCTCGACCCCTC 143
|||||
QY 1479 tcgaactacacgcagcagaggaataattctcccccagcagactgacgatactgggcca 1538
Db 142 TCGAAACTACACGGCAGAGGAGAAATCTTCGCCACGCGACTGATCGCATCTGGGCCAA 83
|||||
QY 1539 ctttcccgacagggatcccaatgagcccgagaccacaaagccccacaatggcccc 1598
Db 82 CTTTGGCCGACAGGGATCCCAATGAGCCCGAGACCCCAAGCCCCCAATGGCCCC 23
|||||
QY 1599 gtacacggcggggctcagcag 1620
Db 22 GTACACGGCGGGGCTCAGCAG 1
|||||

RESULT 10
BE466391/c 562 bp mRNA linear EST 27-JUL-2000
LOCUS hz21b04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3208591 3'
DEFINITION similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN
);.contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION BE466391
VERSION BE466391.1 GI:9512253
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 562)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 469.

FEATURES

source
 1. .562
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3208591"
 /clone_lib="NCI_CGAP_SC6"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 104 a 184 c 173 g 97 t 4 others
 ORIGIN
 Query Match 25.5%; Score 440; DB 10; Length 562;
 Best Local Similarity 99.8%; Pred. No. 1.4e-195;
 Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1137 agacaacagctctctatcagccgagcttcctgcccggggtgcgggtcggggtcc 1196
 |||||
 Db 491 AGACAACGAGCTCTCATCAGCCGGCCGAGTTCCTGGCCGGGTGCGGGTCC 432
 QY 1197 ccaggtaagtacactgagcagcagctggttcctgcattacacagactgctgctcc 1256
 |||||
 Db 431 NCAGGTAAAGTACCTGGCAGCCAGGCTGTGCTGCATTACACAGCTGGTGCATCC 372
 QY 1257 cgaggaccggcagcgcctgaggagccctgagcgtatgtgtggcgaccacacatgtcgt 1316
 |||||
 Db 371 CGAGGACCCGGCAGCCTGAGGAGGCGCCTGAGCGATGTGTGGGCGACCAATGTCTG 312
 QY 1317 gtgcccgtgcccagcgtggtggcgactggtgcccagggtgcccgggtctacgccta 1376
 |||||
 Db 311 GTGCCCGTGGCCAGCTGGCTGGCGACTGGCTGCCAGGGTGCCTGGGTCTACGCCTA 252
 QY 1377 cgtctttgaacacgcgtgctccacgtctctctgcccctgtgcatggggtgccccacgg 1436
 |||||
 Db 251 CGTCTTTGAACACCGGCTCTCCACAGCTCTCTGCGCCCTGTGTGGGTGCGCCACGG 192
 QY 1437 ctacgagatcagttctatctttgggatcccccctggagcccccctcgaacctacacgcgaga 1496
 |||||
 Db 191 CTACGAGATCGAGTTTCATCTTTGGGATCCCTCGGACCCCTCTCGAAACTACACGGCAGA 132
 QY 1497 ggagaaaaatttcgcccagcagctgatgcgatactgggcccacatttgcggcgacagggga 1556
 |||||
 Db 131 GGAGAAAACTTCGCCACGAGCTGATGCGATATCTGGGCCAACTTTGCCGCCACAGGGGA 72
 QY 1557 tcccaatgagccccgagaccccaagccccacacatggccccctacacgcgggggtccta 1616
 |||||
 Db 71 TCCCATGAGCCCGAGACCCCAAGGCCCCCAACATGGCCCCCGCTACACGGCGGGGCTCA 12
 QY 1617 gcagtagcgtta 1627
 |||||
 Db 11 GCAGTAGCTTA 1

RESULT 11

AI190022/c
 LOCUS
 DEFINITION

AI190022 822 bp mRNA linear EST 28-OCT-1998
 gd35e12.x1 Soares fetal_heart_NbHH19W Homo sapiens cDNA clone
 IMAGE:1731502 3' similar to gb:M55040 ACETYLCHOLINESTERASE
 PRECURSOR (HUMAN); mRNA sequence.

ACCESSION

AI190022
 VERSION
 KEYWORDS
 SOURCE

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 822)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1019 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 469.

FEATURES

source
 1. .822
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="1731502"
 /clone_lib="Soares_fetal_heart_NbHH19W"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACCAATCTGAAGTGGAGCGCCGATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."

BASE COUNT

151 a 271 c 245 g 150 t 5 others
 ORIGIN

Query Match

Best Local Similarity 25.4%; Score 438; DB 9; Length 822;
 Matches 538; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1106 ttctggtttacggggccccaggttcagcaagaacacagctctctcatcagccggccg 1165
 |||||
 Db 540 TTCTGTTTACGGGGCCCCAGGCTTCAGCAAGACACAGAGTCTCTCATCAGCGGGCCG 481
 QY 1166 agttcctgcccgggtgctgggttcgccaggttaagtacgtgacgtgagcagagctg 1225
 |||||
 Db 480 AGTTCCTGGCCGGGTGCGGGTTCGCCAGGTAAAGTGAACCTGGACCGCAGGCTG 421
 QY 1226 tggctctgcattacacagactgctgcatcccgaggaccggcgacgcctgagggagccc 1285
 |||||
 Db 420 TGGTCTGCTTACACAGACTGGCTGCATCCGAGGAGCCCGCAGCGCTGAGGAGGCC 361
 QY 1286 tggagcatgtggtggcgaccacaatgtctgtgccccgtggccccagctggctggcgac 1345
 |||||
 Db 360 TGAGCGATGTGGTGGCGACCAATGTCTGTGCNCCGTGGCCACGCTGGCTGGCGAC 301
 QY 1346 tggctcccaggtgcccgggtctcagcctacgtctttgaacaccgtgtcttcacgctct 1405
 |||||


```

Db 300 TGGCTGCCAGGGTGCCCGGCTACGCTACGCTCTTTGAACACCGTGCTTCCACGCTCT 241
QY 1406 cctgcccctctgagtgaggggtgccccaggtacagatcagtgatccttttggatcc 1465
Db 240 CCTGCCCTCTGGATGGGGTGCCTCAGGCTAGGATCAGATTCATCTTTGGGATCC 181
QY 1466 ccttgagccctctgaaactacacggcagagagagaaatcttccacgacgactgatgc 1525
Db 180 CCTTGAGCCCTCTCGAACAATACAGGAGAGAGAGAAATCTTCGCCACGACGTGATGC 121
QY 1526 gatactgggccaactttgcccgcacaggggattcccaatgagccccagagacccaagggccc 1585
Db 120 GATACTGGCCCAACTTGGCCGCACAGGGATCCCAATGAGCCCGAGACCCCAAGGCC 61
QY 1586 cacaatggcccccgtaacagcgggggctcagcagtaagttagttgacctgcggccgc 1645
Db 60 CACAATGGCCCCGTACACGGCGGGGCTCAGCAGTACGTTAGTCTGGACCTGCGGCCGC 1

RESULT 12
BI667712 738 bp mRNA linear EST 12-SEP-2001
LOCUS 603293067F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312273 5',
DEFINITION mRNA sequence.
ACCESSION BI667712
VERSION BI667712.1 GI:15581945
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 738)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraiki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11791 row: f column: 18
High quality sequence stop: 736.
FEATURES
Location/Qualifiers
1..738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312273"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 107 a 231 c 248 g 152 t
ORIGIN
Query Match 24.8%; Score 428; DB 10; Length 738;
Best Local Similarity 99.6%; Pred. No. 6.5e-190;
Matches 528; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctgtgcacacgcttccctggttccccactctctctc 60

```

```

Db 62 ATGAGCCCCCAGTGTCTGTGCACACGCTTCCCTGGCTTCCCACATCTCTTCCTC 121
QY 61 ctctcttgctcctctgggtgagagagtgagggtgagggccggagagatcagagctgctg 120
Db 122 CTGCTGTGCTCCTGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 181
QY 121 gtgacgggtgctggtggggccgggtcggggcatctgcctgaagacccccggggccctgtc 180
Db 182 GTGACGGGTGCTGGGGCCGGCTTGGGGGCTTGGCTGAAGACCCCGGGGCCCTGTC 241
QY 181 tctgttcttctgggcatcccttctcgagagcaccatctgggacccttgcttcttctgcca 240
Db 242 TCTGCTTTCTCTGGGATCCCTTTTGGGAGCACCACCATGGGACCCGCTGCTTCTTGCCA 301
QY 241 ccggagcccaacagccttgggtcagggtggttagacgtcacaccttccagagtgtctgc 300
Db 302 CGGAGGCCCAAGCAGCTTGGTCAGGGGTGGTAGAGCTTCAACCTTCCAGAGTGTCTGC 361
QY 301 taccaatatgtggacaccctataccocaggttttggagggccacgagatgtggaaacccaac 360
Db 362 TACCAATATGTGCACACCCCTATACCAGGTTTGTAGGGCACCAGAGATGTGGAACCCCAAC 421
QY 361 cgtgagctgagcagagactcctgtacctcaacgttgtgacaccatacccccgccctaca 420
Db 422 CGTGAGCTGAGCGAGACTGCCTGTACCTCAACGCTGTGACACCATACCCCGGGCCTACA 481
QY 421 tccccacccctctcctctgtggtatctatgggggtgggtctctacagtggggacctctcc 480
Db 482 TCCCCACCCCTGTCTATCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCTCC 541
QY 481 ttggacgtgtacagtgccgctctctgtgtacagggccgagagagactgtct 530
Db 542 TTGGACGTGTAGATGGCGCTTCTTGGTACAGGGCCGAGAGGACTGTGCT 591

RESULT 13
BI667712 478 bp mRNA linear EST 24-SEP-1998
LOCUS Ox74d09.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1662065
DEFINITION 3' similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);,
mRNA sequence.
ACCESSION AI082805
VERSION AI082805.1 GI:3417781
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 478)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 682 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 463.
FEATURES
Location/Qualifiers
1..478
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1662065"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three

```


REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

1 (bases 1 to 702)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11609 row: g column: 01
 High quality sequence stop: 702.

FEATURES

SOURCE

Accession numbers

```
/organism="Homo sapiens"  
/db_xref="caxon:9606"  
/clone_image="5241624"  
/clone_lib="NIH_MGC_121"  
/lab_host="DH10B"  
/note="Organ: brain; Vector  
Site 2: EcoRV (destroyed);  
fetal brains, female age 20  
and male age 26 weeks. Libr-  
directionally cloned (EcoRV  
cloning). Average insert s-  
ize 0.7-3 kb. Library is nor-  
mal-length clones and was  
(Invitrogen). Research Gen-  
this is a NIH_MGC Library."
a 240 c 226 g 137 t
```

	BASE COUNT	ORIGIN
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0
34	0	0
35	0	0
36	0	0
37	0	0
38	0	0
39	0	0
40	0	0
41	0	0
42	0	0
43	0	0
44	0	0
45	0	0
46	0	0
47	0	0
48	0	0
49	0	0
50	0	0
51	0	0
52	0	0
53	0	0
54	0	0
55	0	0
56	0	0
57	0	0
58	0	0
59	0	0
60	0	0
61	0	0
62	0	0
63	0	0
64	0	0
65	0	0
66	0	0
67	0	0
68	0	0
69	0	0
70	0	0
71	0	0
72	0	0
73	0	0
74	0	0
75	0	0
76	0	0
77	0	0
78	0	0
79	0	0
80	0	0
81	0	0
82	0	0
83	0	0
84	0	0
85	0	0
86	0	0
87	0	0
88	0	0
89	0	0
90	0	0
91	0	0
92	0	0
93	0	0
94	0	0
95	0	0
96	0	0
97	0	0
98	0	0
99	0	0
100	0	0

Query Match	24.2%	Score 417	DB 10	Length 702
Best Local Similarity	99.6%	Pred. No. 9.4e-185		
Matches 517	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy 1	atgaagccccgcagtgctgtcacacgccttccctggcttccccactcttctcctc	60		
Db				
87	ATGAGGCCCCCGCAGTGTCTGTCTGCACACGCGCTTCCTTGGCTTCCCCACTCCTTTCCTC	146		
Qy 61	ctcctctggctcccttgggtggagggagtggggctgagggccggggaggaagcagagctgctg	120		
Db				
147	CTCCTCTGGCTCTTGGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	206		
Qy 121	gtacagtgctggggggccgggctgcggggcattcgctgaagacccccgggggcccctgtc	180		
Db				
207	GTGACGGTGGTGGGGCGGGCTGCGGGGCATTTCGCTGAAGACCCCGGGGGCCCTGTC	266		
Qy 181	tctgctttctgggcattccccctttgcggagccaccctatgggacccctgcgtttctgcga	240		
Db				
267	TCTGCTTTTCTTGGGCATCCCTTTTTCGGAGGACCCCATGGGACCCGCTCGTTCCTTGCCA	326		
Qy 241	ccggagcccaagcagcttggtcaggggtgttagacgtacaaccttccagagtgctgc	300		
Db				
327	CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGTGACGGCTAACCTTCCAGAGTGTCTGCG	386		
Qy 301	taccaatgtggacacccctatacccaggttttgagggacccgagatgtgaaaccccaac	360		
Db				
387	TACCAATATGTGGACACCCCTATACCCAGGTTTTCAGGGCACCCGAGATGTGGAACCCCAAC	446		
Qy 361	cgtagctgagcaggagactgcctatacctaacgtgtgacacataccccccggcctaca	420		
Db				
447	CGTAGCTGAGCGAGGACTGCCTGTACTTCAACGTGTGGACACCATATCCCGGACTACA	506		
Qy 421	tccccaccctctccctcgtcttgatctatgggggtggcttctcagtggggcccctccc	480		
Db				
507	TCCCCAACCCCTGTCCTCGCTGTGATCTATGGGGGTGGCTTCTTACAGTGGGGCCTCCCTCC	566		
Qy 481	ttggagcgtgtacgatggccgctcttcttgtagcaggcccgag	519		

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:43:50 ; Search time 154.73 Seconds
(without alignments)
2738.433 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgagggcccccagtgctctt.....aattgctcagcgccacctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1722	99.8	1845	1	US-07-732-962A-1
2	1722	99.8	1845	5	PCT-US92-06106-1
3	1722	99.8	2256	2	US-08-318-826A-5
4	1722	99.8	2256	2	US-08-370-156-1
5	1722	99.8	2256	3	US-08-814-095-1
6	1722	99.8	3016	2	US-08-318-826A-7
7	1722	99.8	3016	2	US-08-370-156-5
8	1722	99.8	3016	3	US-08-814-095-5
9	1722	99.8	3096	2	US-08-318-826A-6
10	1722	99.8	3096	2	US-08-370-156-3
11	1722	99.8	3096	3	US-08-814-095-3
12	1070	62.0	3506	3	US-08-814-095-7
13	35	2.0	35	2	US-08-370-156-13
14	35	2.0	35	2	US-08-370-156-23
15	35	2.0	35	4	US-08-875-710-4
16	20	1.2	20	3	US-08-990-065-4
17	20	1.2	20	3	US-08-990-065-5
18	20	1.2	20	3	US-08-990-065-7
19	19	1.1	523	4	US-09-437-054A-5
20	19	1.1	3159	4	US-09-437-054A-7
21	19	1.1	246240	2	US-08-724-394A-20
22	19	1.1	246240	2	US-08-724-394A-21
23	19	1.1	246240	2	US-08-724-394A-22
24	18	1.0	300	4	US-09-135-994-3
25	18	1.0	477	2	US-09-109-266-17
26	18	1.0	719	3	US-08-370-223-4
27	18	1.0	940	1	US-08-353-400-24

28	18	1.0	943	2	US-08-303-569B-4	Sequence 4, Appl
29	18	1.0	943	2	US-08-116-247-4	Sequence 4, Appl
30	18	1.0	1243	4	US-09-103-875-16	Sequence 16, Appl
c 31	18	1.0	1276	4	US-09-177-325-2	Sequence 2, Appl
c 32	18	1.0	1276	4	US-09-411-812A-2	Sequence 2, Appl
c 33	18	1.0	1276	4	US-09-590-113-2	Sequence 2, Appl
34	18	1.0	1701	3	US-09-264-737-3	Sequence 3, Appl
35	18	1.0	2184	1	US-08-445-050-8	Sequence 8, Appl
36	18	1.0	2184	1	US-08-204-691-8	Sequence 8, Appl
c 37	18	1.0	2301	1	US-08-306-691B-23	Sequence 23, Appl
c 38	18	1.0	2301	5	PCT-US93-06251-78	Sequence 78, Appl
39	18	1.0	2428	1	US-08-445-050-1	Sequence 1, Appl
40	18	1.0	2428	1	US-08-204-691-1	Sequence 1, Appl
41	18	1.0	2487	3	US-08-370-223-12	Sequence 12, Appl
42	18	1.0	3018	1	US-08-347-718B-3	Sequence 3, Appl
43	18	1.0	3018	1	US-08-482-262-3	Sequence 3, Appl
44	18	1.0	3018	6	5200183-1	Patent No. 5200183
c 45	18	1.0	3123	1	US-08-094-889-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-732-962A-1
; Sequence 1, Application US/07732962A
; Patent No. 5248604
; GENERAL INFORMATION:
; APPLICANT: Fischer, Melr
; TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
; TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07732.962A
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
US-07-732-962A-1

Query Match 99.8%; Score 1722; DB 1; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atagggccccgcagtgctgtgtgacacacgcttccctccctggtcttcccccaactcctctctc 60
DB 1 ATGAGGCCCCCGAGTGTCTGTCACACGCTTCCCTGCTTCCCGACCTCTCTCTCTC 60
QY 61 ctctcttggtctctggtgagagtgagggctgagggccgagagatgacagctgctg 120
DB 61 CTCTCTGCTCTCTGCTGAGAGTGGGGCTGAGGGCCGGAGGATGACAGAGCTGCTG 120
QY 121 gtacaggtgctggtgggcccgcctgagggcattcgccctgaagacccccggggccctgtc 180
DB 121 GTGACGGTGTGGGGCCGCTGCGGGGCTTCCGCTGAAGACCCCGGGGGCCCTGTC 180
QY 181 tctgcttctctgggacccctcttgggagccacccatgggaacccogtcttcttgcca 240
DB 181 TCTGCTTCTCTGGGCATCTCCCTTTGGGAGCCACCCATGGGAGCCCGCTCTTCTGCCA 240
QY 241 ccgagcccaagcagctgtgtgaggggtgtagagctacaaccttccagagtctgc 300
DB 241 CCGAGGCCCAAGCAGCTTGGTCAGGGGTGGTAGACCTACAACCTTCCAGAGTGTCTGC 300
QY 301 taccaatatgtgacacccctatacccagggttttgaggccaccgagatgtgaaaccccaac 360
DB 301 TACCAATATGTGACACCCCTATACCCAGGTTTGAAGGCCACCGAGATGTGAACCCCAAC 360
QY 361 cgtgagctgagcagactcctgtacctcaacgctgagacacacataccccccgacctaca 420
DB 361 CGTGAGCTGAGCAGGACTGCTCTGCTACCTCAACGTGTGGACACATACCCCGCGCTACA 420
QY 421 tcccccaacctctcctcgtctgactataggggtgagcttctacagtgaggctctccc 480
DB 421 TCCCCCACCCCTGCTCTGCTGATCTATGGGGGTGGCTTACAGTGGGGGCTCTCTCC 480
QY 481 ttggacgtgtacgatggccgctctcttggtacagggccgagaggactgtgctggtgctcatg 540
DB 481 TTGGACGTGTACGATGGCCGCTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 540
QY 541 aactacaggtgagagccttctgctctcctgcccctgcccggggagcagagagcccgggc 600
DB 541 AACTACCGGTTGGAGGCTTGGCTTCTCTGGCCCTGCCCCGGGAGCCGAGAGGCCCGGGC 600
QY 601 aatgtgggtctctgatacagagctgagccctcagtgaggtgagagaaactgtggaagcc 660
DB 601 AATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGAGGAGAACGTGGCAGCC 660
QY 661 ttcgggggtgacccgacacatcagtgacgtgtttgggagagcgcggagccgcctcggtg 720
DB 661 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGGAGCGCTCGGTG 720
QY 721 ggcagcactgtgtccccgcagccagcgggctgttccacagggccgctgtgtagagc 780
DB 721 GGATGACCTGTGTCCCCCGCCAGCCGGGGCTGTGTTCCACAGGGCCGTGTCGACAGC 780
QY 781 ggtgcccccaatggaacctggtggccaggtgggcatgagagagcccgctcagggccacg 840
DB 781 GGTGCCCCCAATGGACCTTGGCCACCGTGGGATGGGAGAGCCCGTCCGAGGGCCAGC 840
QY 841 cagctggccccacttbtgggctgtctccagggcggcactggtgggaatgacacagactg 900
DB 841 CAGCTGGCCCCACCTTGTGGCTGTCTCTCCAGGGCGCACTGGTGGGAATGACACAGAGCTG 900
QY 901 gtacgctgcttcggacagcagcagcagctcctggtgaaccacgaatggacgctgctg 960
DB 901 GTAGCTGCTTCCGACACGACGAGCGAGGCTCTGTTGAACCAACCAATGGCAGCTGCTG 960
QY 961 cctcaagaagcgtcttccgggttctctctgctgctggtgtagatgagagcttctcagt 1020
DB 961 CCTCAAGAAGCGCTTCTCCGTTCTCTCTGCTGCTGTGTAGATGGAGACTTCTCAGT 1020
QY 1021 gacacccagagccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
DB 1021 GACACCCAGAGGCCCTCATCAACGGGGAGACTTCCACGGCTGCAGGTGCTGGTGGT 1080
QY 1081 gtggtgagagatgagggctgtatttctgttttagggggcccgagcttcagcaagac 1140
```

```
DB 1081 GTGGTGAAGGATGAGGCTCGTATTTTCTGTTTACGGGGCCCGAGCTTTCAGCAAGAC 1140
QY 1141 aacgagctctctcatcagccgggcccaggttctgcccggggtcggggtctccccag 1200
DB 1141 AACGAGTCTCTCATCAGCCGGGCGGAGTTCTTGCCCGGGGTGGGGTCCGGTTCGCCAG 1200
QY 1201 gtaagtacacctggcagccgagggctgtggtcctgcatatacacagactggtgcacccag 1260
DB 1201 GTAAGTGACCTGCGAGCCGAGGCTGTGCTCTGCTTACACAGACTGGCTGCATCCCGAG 1260
QY 1261 gacccggcagcctgagggggccctgagcgtatggtggggcagacacaatgtcgtgtgc 1320
DB 1261 GACCCGGCAGCCCTGAGGGAGGCCCTGAGCGATGTGTGGGGCAGCAACAATGTCTGTGC 1320
QY 1321 ccgctgccccagctggtggtgagctggtgcccaggttcccgggtctctacgctcagtc 1380
DB 1321 CCCGTGGCCCAAGCTGGCTGGCGACTGGCTGCCAGGGTGCCCGGGTCTACGGCTACGTC 1380
QY 1381 tttagaacacctgtcttccacgctctcctggtcccctgtggatgggggtgcccacggctac 1440
DB 1381 TTTGAACACCGTGTCTCACGCTCTCTCTGCCCCCTGTGGATGGGGTGGCCACGGCTAC 1440
QY 1441 gatatcaggttcatcttggggtacccctggaacccctctcgaactacacagggagag 1500
DB 1441 GAGATCGAGTTCATCTTTGGGATCCCCCTGACCCCTCTCGAAACTACACGGCAGAGAG 1500
QY 1501 aaaaatctcccccagcagctgctgctgactgctggtggtggtggtggtggtggtggtggt 1560
DB 1501 AAAATCTTCCCCCAGCAGCTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560
QY 1561 aatgagccccgagaccccaaggcccccaaatggtcccctgtacacggtggtggtcagcag 1620
DB 1561 AATGAGCCCCGAGACCCCAAGGCCCAACAATGGCCCCGTACACGGGGGGGCTCAGCAG 1620
QY 1621 taagttagctgacactgagccgctgaggtgaggtgagggggggtgaggtgaggtgaggtgag 1680
DB 1621 TAGCTTAGTCTGACCTTGGCGCGCTGGAGGTGGCGGGGGGCTGGCGGCCCGCCAGGCGTGC 1680
QY 1681 gcttcttggaacgcttctcccaaatgctcagcgccacc 1722
DB 1681 GCCTTCTGGAACCGCTTCTCTCCCAAAATTCCTCAGCGCCACC 1722
```

RESULT 2

```
PCT-US92-06106-1
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06106
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
PCT-US92-06106-1

Query Match 99.8%; Score 1722; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgaggccccagcagtgctgtgcacacgccttccctggcttccccactccttctctc 60
Db 1 ATGAGGCCCCGCGAGTGTGTCGCACACGGCTTCCTCGGCTTCCCCACTCCTTCTCCTC 60

Qy 61 ctctctgctcctggtgaggagctgggctgaggccggagccggagatgcagagctgtg 120
Db 61 CTCCTCTGGCTTCCCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCGAGAGTGTGTG 120

Qy 121 gtgacgggtgctggggcgccgctgcgggcatctccctgaagacccccggggccctgtc 180
Db 121 GTGACGGGTGCTGGGGCGCGGCTGCGGGGCATTCGCTTGAAGACCCCGGGGGCCCTGTC 180

Qy 181 tctgtttctcggggaaccctttctgagagccacccaatggaccctgtcgtttcttgc 240
Db 181 TCTGTTTCTCGGGCATCCCTTTTCGGAGCCACCCATGGGACCCTCGCTTCTTGCCA 240

Qy 241 ccggagcccaagcagcctgttcagggtgttagcgtacaaacctccagagtgctgc 300
Db 241 CCGGAGCCCAAGCAGCCTTGTTCAGGGGTGTGAGCGTACAACTTCCAGAGTGTCTGC 300

Qy 301 taccaatatgtggacacccctataccagggttttgaggccacccagatgtggaaaccccaac 360
Db 301 TACCAATATGTGGACACCCCTATACCAGGTTTGTAGGGCCACCGAGATGTGGAACCCCAAC 360

Qy 361 cgtgagctgagcagagactgctgtacctcaacgtgtggacacccataccccggccta 420
Db 361 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTTGTGGACACCATACCCCGCGCTACA 420

Qy 421 tccccacccctgtcctcgtctggtatctatggggctggtcttctacagtggggcctctcc 480
Db 421 TCCCCACCCCTGTCTCGTCTGGAATCTATGGGGTGTGCTTCTACAGTGGGGCTTCTCC 480

Qy 481 ttggagctgtacgatggccgtcttctgttacaggccgagagactgtgctggtgtccatg 540
Db 481 TTGGAGCTGTACGATGGCCGTCTTGTGTACAGGCCGAGAGACTGTGCTGTGTCCATG 540

Qy 541 aactaccgggtggagaccttggcttctcctgcccctgcggggagccgagagccccgggc 600
Db 541 AACTACCGGGTGGAGACCTTTGGCTTCTGCCCCCTGCGGGGAGCCGAGAGGCCCGCGGC 600

Qy 601 aatgtgggtctccttgatcagaggtgcccctgagtgagtgagtgaggaacagtgagacc 660
Db 601 AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCGAGTGGGTGCAGGAGAACGTGGCAGCC 660

Qy 661 ttgggggtgaccagacatcagtgacgtgttttggggagagcgcgaggccctcgggtg 720
Db 661 TTGGGGGTGACCCACATCAGTGACGTGTTTGGGGAGAGCGCGGAGCCCTCGGTG 720

Qy 721 ggcagtcacactgtctcctcccgcccgccgggcccctgtttccacagggccgtgtcagagc 780
Db 721 GGCATGCACCTGCTCTCCCGCCCGCCAGCGGGGCCCTGTTCCACAGGGCCGTGTGCGAGAC 780

Qy 781 ggtgcccccaatggaccctgggcccagcgtgggcacatggggagagggccgctgcagggccacg 840

Db 781 GGTGCCCAATGGACCCCTGGGCCACGGTGGCATGGGAGAGGCCGCTGCGAGGGCCACG 840

Qy 841 cagctggccacacctgtgtggctgtcctccaggcggcactggtgggaaatgacacagagctg 900
Db 841 CAGCTGGCCCACTTGTGGCTGTCTCCAGGGCGGCACTGGTGGGAATCACACAGAGCTG 900

Qy 901 gtactgctcctcgacacgaccagcaggtctctgtgtaaacacagaaatggcacctgtctg 960
Db 901 GTAGCTTGCTTCGGACACGACGACGAGGCTCCTGGTGAACCAAGAAATGGCACGCTGCTG 960

Qy 961 cctcaagaaagcgtcttcgggttctccttcgtcctctgtgtagatgagagacacttcctc 1020
Db 961 CCTCAAGAAAGCGTCTTCGGGTTCCTTCGTGCTCTGTGTAGATGAGAGACTTCTCTCAGT 1020

Qy 1021 gacacccagagggccctcatcaacgcgggagacttccacggctccacggctgcaggtgctgg 1080
Db 1021 GACACCCAGAGGCCCTCATCAACGGGGAGACTTCCACGGGCTGCAGGTGTGTGGGT 1080

Qy 1081 gtggtgaaggatgagggctgtctattctgttctcgttcaaggggccccaggcttcagcaagac 1140
Db 1081 GTGGTGAAGGATGAGGGCTGCTATTCTGTGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1140

Qy 1141 aacgagctctcatcagcggcgccaggttctcggcggggtgcggggtcggggttcccccag 1200
Db 1141 AACGAGTCTCTCATCAGCGGGGCGGAGTTCTTGCCCGGGGTGCGGGTTCGCCAG 1200

Qy 1201 gtaagtgaactggcagcggaggtgtggtcctgcattacacagactggctgcatcccgag 1260
Db 1201 GTAAGTGACTGGCAGCCGAGGCTGTGTGCTTGCATTACACAGACTGGGTGTGATCCCGAG 1260

Qy 1261 gaccggcaacgctcagggaggccctgagcgtgtgtgtggcgacaacaatgtcgtgtgc 1320
Db 1261 GACCGGCGCAGCGCTGAGGGAGGCCCTGAGCGATGTGTGGGGAGCCACCAATGTCTGTGTGC 1320

Qy 1321 ccgtggcccaactgctggcgactggctgccaggggtgccgggtctacgctacgtc 1380
Db 1321 CCGTGGCCCACTGCTGGCGGACTGGCTGCCAGGGGTGCCCGGGTCTACGGCTTACGTC 1380

Qy 1381 ttgaaacacgctgtcttccacgctcctcgtggccctgtggtggtgggtgccccacggctac 1440
Db 1381 TTTGAACACCGTGTCTCCACGCTCTCTGGCCCCCTGTGATGGGGTGGCCCCACGGCTAC 1440

Qy 1441 gagatcgagttcaatttgggataccccctggaccctctcgaactacacgcagaggag 1500
Db 1441 GAGATCGAGTTTCACTTTGGGATCCCTTGGACCCCTCTCGAAACTTACACGGCAGAGGAG 1500

Qy 1501 aaaaatttcggccagcagactgatcgatactggcacaactttgccgcacaggggagtc 1560
Db 1501 AAAATCTTCCGCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560

Qy 1561 aatgagccccgagaccccaaggccccacaaatggccccctacacggcggggggtcagcag 1620
Db 1561 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCTACACGGCGGGGCTCAGCAG 1620

Qy 1621 tacgttagctcagacactgcggcgtgaggtgcggcggggggtgcgcgccagggcctgc 1680
Db 1621 TACGTTAGCTCAGACCTGCGGCCGCTGGAGGTCGCGGGGGGCTGCGCGCCCGCCAGGCGCTGC 1680

Qy 1681 gcctctggaacgcttctccccaaaattgctcagcgccacc 1722
Db 1681 GCCTTCTGGAAACCGCTTCTCCCCAAATTGTCTACGGCCACC 1722

RESULT 3
US-08-318-826A-5
; Sequence 5, Application US/08318026A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense

;; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
;; TITLE OF INVENTION: Containing Them
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kohn & Associates
;; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: US
;; ZIP: 48334

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/318,826A
;; FILING DATE:
;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: 2391.00001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 539-5050
;; TELEFAX: (248) 539-5055
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA to mRNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
;; OTHER INFORMATION: 3, 4 and 6"

US-08-318-826A-5

Query Match 99.8%; Score 1722; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagccccccagtgctgtcgcacagccttccctgagtcctccacactccttctc 60
|||||
Db 160 ATGAGCCCCCGGAGTGCTGCTGCGACACGCCCTCCCTGCTTCCCGACCTCTCCCTC 219
QY 61 ctctctggtctcctggtgagtgagtggtggtggtgagggccgggagtgagagtgctg 120
|||||
Db 220 CTCCTCTGGCTCTCTGGTGAGGAGTGTTGGGGCTGAGGGCCGGGAGGATGCAGAGTGTG 279
QY 121 gtagcagtgctggtggggccgctgctgctgctgctgctgctgctgctgctgctgctg 180
|||||
Db 280 GTACAGGGTGGGTGGGGCCGGCTGCGGGGCAATTCGCTGAAGACCCCGGGGGCCCTGTC 339
QY 181 tctgcttctcctgggcatcccttctgagagccaccatggaccctgctgcttctgcca 240
|||||
Db 340 TCTGCTTCTTGGGCATCCCTTTTGGGAGCCACCCATGGGACCCCGCTCTCTTCTGCCA 399
QY 241 ccggagcccaagcagccttggtcaggggtggtgagcgtacacaccccttccagagtgtgc 300
|||||
Db 400 CCGAGGCCCAAGCAGCTTGGTCAGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 459
QY 301 taccataatgtgacacccctataccacagttttgagggccacccagagatgtggaacccaac 360
|||||
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTTTGGAGGCCACCCAGAGATGTGAAACCCCAAC 519
QY 361 cgtgagctagcagagactgcttaccctacacgtgtagacacatacccccgggctaca 420
|||||

Db 520 CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGCTGTGGACACCAATACCCCCCGCCTACA 579
QY 421 tccccacacccctgtcctcgtctggatctatgggggtgcttctacagtggggctcctcc 480
|||||
Db 580 TCCCCACACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCC 639
QY 481 ttggacgtgtacgatggcctcttcttggtacagccagagagactgtgctggttccatg 540
|||||
Db 640 TTGGAGCTGTACGATGCCCTTCTTGGTACAGGCCGAGAGGACTGTCTGGTGTCCATG 699
QY 541 aactacgggtggagcccttctgcttctcctcctcctcctcctcctcctcctcctcctcctc 600
|||||
Db 700 AACTACCGGGTGGAGCCTTTGGCTTCTTGGCCCTGTCGGGGAGCCGAGAGGCCCGGGC 759
QY 601 aatgtgggtcctcctggatcagagcctgcccctgcagtggtgagagagacgtggcagcc 660
|||||
Db 760 AATGTGGGTCTCTCTGGATCAGAGCCTGGCCCTGCAGTGGGTGCAGGAGAACTGGCAGCC 819
QY 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcggggagccgctcctcctg 720
|||||
Db 820 TTCCGGGGTGCACCGACATCAGTACGCTGTCTTGGGAGAGCGCGGAGCGCGCTCGGTG 879
QY 721 ggcacacactgctgtcccgcacccagcgggctgttccacagggccgtgctgacagc 780
|||||
Db 880 GGATGCACCTGTCTCCCGCCAGCCGGGCTGTTCACAGGGCCGTGTCTGTCAGAGC 939
QY 781 ggtgcccccaatggacccctgggacagtggtgagagagcggcctgcagggccacg 840
|||||
Db 940 GGTGCCCCCAATGACCTTGGCCACCGTGGCATGGGAGAGCCCGCTCCAGGGCCACG 999
QY 841 cagctggccacacttgggctgtcctccagggcggcactggtgggaatgacacagactg 900
|||||
Db 1000 CAGCTGGCCCACTTGTGGCTGTCTTCCAGGGGCACTGTGTGGGAATGACACAGAGCTG 1059
QY 901 gtagcctgcttcggacacagcagcagctcctggtgagagagcggcctgacacagcgtg 960
|||||
Db 1060 GTAGCTGCTTCCGACACACGACGAGGCTCTGTGTGAACCAACGAATGGCAGTGTG 1119
QY 961 cctcaagaaagcgtcttccgggttctcctcgtgctgtgtagatgagacttccctcagt 1020
|||||
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCCTCTGCTGCCCTGTGTAGATGGAGACTTCTCAGT 1179
QY 1021 gacacccagagccctcatcaacgcgggagacttccacggcctgaggtgcgtgggtg 1080
|||||
Db 1180 GACACCCACAGAGCCCTCATCAACGGGGAGACTTCCACGGCCTGCGAGTGTGGTGGT 1239
QY 1081 gtggtgaagatgagggctgtatttctggtttacggggcccccagccttcagcaagac 1140
|||||
Db 1240 GTGCTGAAGATGAGGGCTCGTATTTCTGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1299
QY 1141 aacgagctctcatcagcggggcgaggttctcctgcccgggtgcgggttcgggttccccag 1200
|||||
Db 1300 AACGAGTCTCTCATCAGCGGGCGGAGTTCCTGCGGGGTGCGGGTTCGGGTTCGCCAG 1359
QY 1201 gtaagtgaacctgacgagcaggtgtggtcctcattacacagactggtcgtatccccag 1260
|||||
Db 1360 GTAAGTGACCTGACCGGAGGTGTGTCTCATTTACACAGACTGGCTGCTATCCCGAG 1419
QY 1261 gacccgcagcctgagggagggccctgagcgtgtggtggcgacacacatgtcgtgtgc 1320
|||||
Db 1420 GACCCGGCAGCCTGAGGGAGGGCCCTGAGGCGATGTGTGGGGCCACCAATGTGCTGTGC 1479
QY 1321 ccggtggccagctggtggtggcactggtgcccaggggtcccgggttctacgcctacgtc 1380
|||||
Db 1480 CCCGTGGCCAGCTGTGCTGGCGACTGGCTGCCAGGGTCCCGGGTCTACGCTACGTC 1539
QY 1381 ttgacacacgctcttcacagctctcctggccctgtggtgggggtgccccacgctac 1440
|||||
Db 1540 TTTGAACACCGCTCTTCCAGCTCTCCTGGCCCTGTGTGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcaggttctcttggatcccccttgacccctctcgaactacacagcagagag 1500
|||||
Db 1600 GAGATCGAGTTCAATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659

QY 1501 aaaaattcttgcggcagcagactgatcgatctgatactggtggccaaatttgcgcgcacaggggatccc 1560
DB 1660 AAAATCTTTGGCCACGAGCTGATGGGCACTTTGGCCGACACAGGGATCCC 1719
QY 1561 aatgagccccagaccccccaagcccccaaatggtcccccgatcacagcgagggtctcagcag 1620
DB 1720 ATAGAGCCCCGAGACCCCAAGCCCCACATATGGCCCCCGTACACGCGGGGCTCAGCAG 1779
QY 1621 tacgttagcttgagactcgggccgctgaggtgaggtgcgggggggctgcgcgcacagcctgc 1680
DB 1780 TAGCTTAGTCTGAGACTCGCGGCGCTGGAGGTGCGCGGGGCTGCGCGCCGAGGCTGC 1839
QY 1681 gaccttggaacccgttctctcccccaattgctcaagcgcacc 1722
DB 1840 GCGTCTTGGAACCGCTTCTCCCCAAATGCTCAGCGCCACC 1881

RESULT 4

US-08-370-156-1

; Sequence 1, Application US/08370156

; Patent No. 5932780

; GENERAL INFORMATION:

; APPLICANT: Soreq, Hermona

; APPLICANT: Zakut, Haim

; APPLICANT: Shani, Moshe

; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR

; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Relasing, Ethington, Barnard & Perry

; STREET: P.O. Box 4390

; CITY: Troy

; STATE: Michigan

; COUNTRY: US

; ZIP: 48099

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/370.156

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Kohn, Kenneth I.

; REGISTRATION NUMBER: 30,955

; REFERENCE/DOCKET NUMBER: P-307 (Mulford)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (810) 689-3500

; TELEFAX: (810) 689-4071

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2256 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-370-156-1

Query Match

Best Local Similarity 99.8%; Score 1722; DB 2; Length 2256;

Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctgtgcacacgccttcttccctggcttccccactcttctcttc 60

DB 160 ATGAGGCCCCGACATGCTGCTGCACACGCGCTTCCCTGGCTTCCCCACTCTCTCTCTC 219

QY 61 ctctcttggtctctgggtgaggtggtggggtgagggcgagggcgagagtgacagagctgtg 120

DB 220 CTCCCTCTGGCTCTGGGTGGAGGAGTGGGGGCTGAGGGCGGGGAGGATGCAGAGCTGTG 279

QY 121 gtagcggtycgtggtggggcggtgcgggggcatctgcctgaagacccccgggggcccctgtc 180
DB 280 GTGACGGTGCgtggggcggtgcggggcatttgcgctgaagacccccgggggcccctgtc 339
QY 181 tctgtttctctgggcatcccttttgcggagccaccatctgggaccccgctgcgtttcttgcca 240
DB 340 TCTGTCTTCTGGGATCCCTTTGCGAGACCCATCGGACCCGCGTGCCTTCTTCTGCGCA 399
QY 241 ccggagcccaagcagccttgggtcagggtggtagacgctcacaccttccagagtgctctgc 300
DB 400 CCGGAGCCCCAACGCTTTGGTCAAGGGTGGTAGACGCTTCAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctataccacaggttttggaggccacgagatgtggaaccccaac 360
DB 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcagagactcctgtacctcaacgtgtgagaccataccccggcgctaca 420
DB 520 CGTGAGCTGAGGAGGACTGCTGTACCTAACGTGTGGACACCATACCCCGCGCTACA 579
QY 421 tccccacccctgtcctcgtctggtatctatggtgggtggtcttctcagtggtgggctctcc 480
DB 580 TCCCCACCCCTGTCTCTGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCTCC 639
QY 481 ttggacgtgtacagatggccgcttctgtgtacagggcgagagactgtgctggtgtccatg 540
DB 640 TTGGACGTGTACGATGGCGCTTCTGTGTACAGGGCGAGAGACTGTGTGTGTCTCATG 699
QY 541 aactaccgggtggagccttggcttctcgtccctgcggggcgagcagagagggcgccggcg 600
DB 700 AACTACCGGGTGGAGCCCTTGGCTTCTGGCCCTGCGGGGAGCGAGAGGCCCGCGGC 759
QY 601 aatgtgggtctcctggtacagaggtggcctgcagtggtgcaggagagactggtggcagcc 660
DB 760 AATGTGGGTCTCTCGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGGAGCC 819
QY 661 ttccgggggtgacccgacatcagtgacgctgttttggggagagcgcgggagccgctcgtg 720
DB 820 TTCCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGAGCCGCTCGGTG 879
QY 721 ggcctgacctgtctcccgcccgagcggtcctgttccacagggcgctgtctgacagac 780
DB 880 GGCATGACCTCTGTCCCGCCCGAGCGGGGCGCTTCCACAGGGCGCTGTCTGCAGAGC 939
QY 781 ggtgcccccaatggaccctggccacggtgggcatgggagagggccgctgcaggggcccag 840
DB 940 GGTGCCCCCAATGGACCTTGGGCCACGGTGGCATGGGAGAGGCCGCTGCAGAGGCCACG 999
QY 841 cagctggcccccacttgtgggtgtctctcagggcgccactggtgggaatgacacagagctg 900
DB 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTTCCAGCGCGCACTTGTGGGAATGACACAGAGCTG 1059
QY 901 gtacgctgcttcgacacgacgagcgaggtccctggtgaaccacgaatggcagctgctg 960
DB 1060 GTAGGCTGCTTCGGACACGACGCGGACGCTGCTGTGTGAACCAAGATGGCAGCTGCTG 1119
QY 961 cctcaagaaagcgtcttccggttctccttctgtgcctgtgtgtagatggagacttctcact 1020
DB 1120 CCTCAAGAAAGCGTCTTCGGTCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
QY 1021 gacacccagagggccctcatcaaacgcgagagacttccacggcctcaggtgctgtggtg 1080
DB 1180 GACACCCGAGAGGCGCTCATCAACGCGGGAGACTTTCACAGGCGCTGCAGGTGTGTGTGT 1239
QY 1081 gtgtgagagatgagggctcgtattttctgttttccggggggccccaggcttccagcaagac 1140
DB 1240 GTGTGTGAAGGATGAGGGCTGCTGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1299
QY 1141 aacgagctctctcatcagccgggcccaggttccctggtgggggtggtggtggtggttccccag 1200
DB 1300 AACGAGTCTCTCATCAGCGGGCGGAGTTCCTGTGGCGGGGTGCGGGGTGCGGGGTCCCCAG 1359
QY 1201 gtaagtgcctggcagcagcagaggtgtgttctgtcattacacagagactggtgctatcccgag 1260

```
|||||
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGGCTCCTGCATTACACACTGGCTGCATCCGGAG 1419
Qy 1261 gacccggacgctgaggggcccctgagcgtatgtgtgggacacacaaatgtctgtgtgc 1320
Db 1420 GACCCGGACACGCTGAGGAGGAGCCCTGAGCGATGTGTGGGGACACACAATGTCTGTGC 1479
Qy 1321 ccsgtggccagctgctggggagctgctgcccaggggtgcccgggtctacgctacgtc 1380
Db 1480 CCCGTGGCCAGCTGGCTGGGGCACTGGCTGCCAGGGTCCCGGGTCTACGCTACGTC 1539
Qy 1381 ttgaaacacgctctccacgctctcctgcccctgtggatgggtgcccacggctac 1440
Db 1540 TTTGAACACCGCTGCTTCCACGCTCTCTGGCCCCCTGTGGATGGGGTGGCCCCACGGCTAC 1599
Qy 1441 gagatgagttcatctttggatccccctggaccctctcgaactacacacgagggag 1500
Db 1600 GAGATCGAGTTTATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
Qy 1501 aaaaattcccgagagactgatactggtgggacaaattttgcccgacagggatccc 1560
Db 1660 AAAATCTTCCCGCAGGACTGATGCGTACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
Qy 1561 aatgagcccgacaccccaagggccacaaatggccccctgacaggggggctcagcag 1620
Db 1720 AATGAGCCCGCAGACCCCAAGGCCCCACAAATGGCCCCCTGACAGGGGGGGCTCAGCAG 1779
Qy 1621 tacgttagtctgacctggcgctggaggtgctgggggggtgctggcgccccagggctgc 1680
Db 1780 TAGCTAGTCTGACCTGCGCGCGCTGGAGGTGGGGGGGGGGCTCGCGCGCCAGGCTGC 1839
Qy 1681 gcctctggaacgctctctcccccattgtctcagagccacc 1722
Db 1840 GCCTTGTGAACGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1881
```

RESULT 5

```
US-08-814-095-1
; Sequence 1, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zukut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "ACHE gene comprising exons
; DESCRIPTION: 2, 3, 4 and 6"
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-814-095-1
```

```
Query Match 99.8%; Score 1722; DB 3; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 atagagccccccagtcgtctgtgcacacgccttccctggcttcccacactccttctcctc 60
Db 160 ATGAGGCCCCCGAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCCACCTCTTCTCCTC 219
Qy 61 ctctctggctctggtggagagtggtgggctgagggccgggagggatgcagagctgctg 120
Db 220 CTCTCTGGCTCTGGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGACAGAGCTGCTG 279
Qy 121 gtacaggtgctggggggccgtgctgggggcttcgctgaagacccccgggggccccctgc 180
Db 280 GTACAGGTGCTGGGGCCGCTGCGGGGCTGCGCTGAAGACCCCCCGGGGGCCCTGTC 339
Qy 181 tctgcttctgggcatcccttctggagagccacccatggagccctgctcttctgcca 240
Db 340 TCTGCTTCTTGGGCACTCCCTTTTGGGAGCCACCCATGGACCCCTGCTCTTCTGCCA 399
Qy 241 ccggagcccaagcagctgtgtgtaggggtgtagacgtacacacattccagagtgctgc 300
Db 400 CCGAGCCCAAGCAGCCTTGTGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
Qy 301 taccataatgtgacacccctatacccaaggttttgaggggacccagagatgtggaaccccaac 360
Db 460 TACCAATATGTGACACCCCTATACCCAGAGTTTGTAGGGCACCCGAGATGTGGNAACCCCAAC 519
Qy 361 cgtgagctgagcagaggaactcctgtactcaacgtgtggacacacacaccccccgcctaca 420
Db 520 CGTGAGCTGAGCGAGGAGTACCTGCTGCTACCTCAACGTGTGACACCATACCCCGCCCTACA 579
Qy 421 tccccacccctgctctgctgtggtatctatgggggtggtctctacagtggggcctctcc 480
Db 580 TCCCCACCCCTGTCTCTGCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC 639
Qy 481 ttgagcgtgtaacatggcgcctctctgttacagccgagagggagactgtgctggttccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGGACTGTGTGTGTGTCCATG 699
Qy 541 aactaccgggtggagcctttgcttctctgcccctgcccctgcccgggagccgagagcccgcc 600
Db 700 AACTACCGGGTGGAGCCTTTGGCTTCTGCCCCCTGGGGGAGCCGAGAGCCCCCGGGC 759
Qy 601 aatgtgggtctcttgatcagagaggtggccctgagtggtgtgaggaagacgtggagacc 660
Db 760 AATGTGGGTCTCTCTGGATCAGAGGCTGCGCCCTGACAGTGGGTGACAGGAACGTGGCAGCC 819
Qy 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcgaggagcgcctcggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTGTGGGAGAGCGCGGAGGCGCTCTCGGTG 879
Qy 721 ggcattgacactgtgtcccccgcagccggggcctgttccacagggccgtgctgtagagc 780
Db 880 GGCATGCACTGTGTGTCCCCCGCCAGCGGGGGCTGTGTCCACAGGGCGCTGTGTGCAAGC 939
Qy 781 ggtgcccccaatggacccctgggcccacggtgggcatgggagagggccctcagagggccacg 840
Db 940 GGTGCCCCCAATGGACCTTGGGCCACGCTGGGCATGGAGAGGGCCCTCTCGAGGGCCACG 999
```

```
Qy 841 cagctggccaccctgtgggctgtcctccagcgagcagctgtggaatgacacagagctg 900
Db 1000 CAGCTGGCCCAACCTTGTGGGCTGTCTCCAGCGGCACTGTGTGGAATGACACAGAGCTG 1059
Qy 901 gtagcctgcttcgacacagcaccagcagctcctggtgaaccacgaatggcagctgctg 960
Db 1060 GTAGCCTGCCCTTCGACACAGCAGCAGCAGCTCCTGTGTGAACCAAGATGACAGTGTCTG 1119
Qy 961 cctcaaaaaagcgttcttcggttctcctctgctgtggttagatggaattctctcagt 1020
Db 1120 CCTCAAGAAACGCTTCTCCGGTCTCTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
Qy 1021 gacacccagagccctcatcaacgagcgagacattccacggcctgacagctgctggtggt 1080
Db 1180 GACACCCAGAGGCCCCCTCATCAACGCGGAGACTTCCACGCGCTGCAGGTGCTGTGTGTGT 1239
Qy 1081 gtgtgnaaggatgaggggtctgtatttttctgtttacggggggccagggcttcagcaaaagac 1140
Db 1240 CTGCTGAAGGATGAGGGCTCTATTTTCTGTGTACGGGGCCCCAGGGCTTCAGCAAAAGAC 1299
Qy 1141 aacgagctctcatcagcggggccagagttcctgcccgggggtgcggggttcgggggttc 1200
Db 1300 AACGAGTCTCTCATCAGCGGGGCGAGTCTGTGCGGGGGTCCGGGTCCGGGTTCGCCAG 1359
Qy 1201 gtaagtgacctggcagcgagctgtgtctctgctcattacacagactgctcatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCGGAGGCTGTGTCTCTGATTAACAGAGCTGTGTCTGATCCCCGAG 1419
Qy 1261 gaccgggcagcctgagggagggcctgagcagatgtggggcgacacaaatgctgtgctg 1320
Db 1420 GACCCGGCACGCTTGAGGAGGCGCTGAGCGATGTGTGTGGGGGACCAACAATGCTGTGTGC 1479
Qy 1321 cccgtggccagctggtggtggcgactggtcccgaggggtgcgggggttcagcctacgtc 1380
Db 1480 CCCGTGGCCCAAGCTGTGCTGGCGAGCTGTGCTGCCAGGGGTGCCGGGTCTACGCCCTACGTC 1539
Qy 1381 ttgaacacagctgtccacgctctcctgcccctgtggtggtggtggtggtggtggtggtggt 1440
Db 1540 TTTGACACAGCTGTCTCCACGCTCTCTCTGGCCCCCTGTGTGTGTGTGTGTGTGTGTGTGT 1599
Qy 1441 gagatcgagttcatttttggggtatccccctggacccctctgaaacctcgaacctcagcagagag 1500
Db 1600 GAGATCGAGTTTCACTTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
Qy 1501 aaatcttcgcccagcagactgagcagatctggtggtggtggtggtggtggtggtggtggtggt 1560
Db 1660 AAAATCTTTCGCCCAAGCGACTGATGCGATCTGTGCGCCAACTTTTGGCCGACACAGGGGATCCC 1719
Qy 1561 aatgagcccgagaccccaagcccaatgagccctgacccctgacccctgacccctgacccctgac 1620
Db 1720 AATGAGCCCCGAGACCCCAAGCCCCCAATGAGCCCCGATACAGCGGGGGGCTCAGCAG 1779
Qy 1621 tacgttagctggaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1680
Db 1780 TAGCTTAGCTGAGCCTGCGCGCGCTGAGAGTGTGCGCGGGGGCTGCGCGCGCCAGGCGCTGC 1839
Qy 1681 gcctcttggaacccgttctctcccaaaatgctcagcgccacc 1722
Db 1840 GCGTCTTGGAAACCGCTTCTCCCAAAATGCTCAGCGCCACC 1881
```

RESULT 6
US-08-318-826A-7
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9

```
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kohn & Associates  
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410  
; CITY: Farmington Hills  
; STATE: Michigan  
; COUNTRY: US  
; ZIP: 48334  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/318,826A  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kohn, Kenneth I.  
; REGISTRATION NUMBER: 30,955  
; REFERENCE/DOCKET NUMBER: 2391.00001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (248) 539-5050  
; TELEFAX: (248) 539-5055  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3016 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 160..2010  
; OTHER INFORMATION: /note= "Splice Variant: Exons 1, 2,  
; OTHER INFORMATION: 3, 4, 5 and 6"  
US-08-318-826A-7
```

```
Query Match 99.8%; Score 1722; DB 2; Length 3016;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 atgagggccccgcagctgtctgtcacacgcttccctggcttccccactctctctc 60  
Db 160 ATGAGGCCCCCGCACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 219  
Qy 61 ctctctggctctctgggtggagagtggtggggctgagggcgaggagatgcagagctgtg 120  
Db 220 CTCCTCTGGCTCCTGGTGGAGAGTGGGGGCTGAGGGCGGGAGGATGCACAGCTGTGTG 279  
Qy 121 gtgacggtgtgtggggcggtgctgggggcttcctcctgaagacccccggggccctgtc 180  
Db 280 GTGACGGTGTGGGGGCGGCTGTGGGGGCAATTCGCTTGAAGACCCCGGGGGCCCTGTGC 339  
Qy 181 tctgcttctctgggcatcccttttgcgagccacccatgggaccccgctcgttctgtcca 240  
Db 340 TCTGTCTTCTTCTGGGCATCCCCCTTTCGGAGGACCCCATCCCTGCTGCTTCTGTGCA 399  
Qy 241 ccggagcccaagcagccttgggtcaggggtggtagacgtctacaacctccagagtgctgc 300  
Db 400 CCGGAGCCCAAGCAGCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459  
Qy 301 taccaatatgtgacacccctatacccccaggtttttgagggcgaccagagatgtggaaccccaac 360  
Db 460 TACCAATATGTGACACCCCTATACCCAGGTTTGTGAGGCGACCGAGATGTGGAACCCCAAC 519  
Qy 361 cgtgagctgagcgagagctgctgtacctcaacgtgtgagacaccataccccggcctaca 420  
Db 520 CGTGAGCTGAGCGAGGACTGCTGTACCTTCAACGTTGTGAGACCAACCATACCCCGGCGCTACA 579
```


Db 220 CTCCTCTGGCTCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGGATGCACAGCTGCTG 279
Qy 121 gtacgggtgcgtggggccgggtgcggggcatttcgacctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGTGGGGCGGGCTGGGGGCAATTCGCCGTGAAGACCCCGGGGGCCCTGTC 339
Qy 181 tctcttctctgggcatccctttgcggagccacccatgggaccccgctgctttctgcca 240
Db 340 TCTGCTTTCTTGGGGATCCCTTTTGGGAGCCACCCTATGGGACCCCGCTGCTTTTGCCA 399
Qy 241 ccggagcccaagcagccttgggtcagggggtggtagcgtctacaacattccagagtctgc 300
Db 400 CGGGAGCCCAAGCAGCCTTGGTCAAGGGTGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
Qy 301 taccaatatgtgaacacctatataccagggttttggaggccacgagatgtgaacccccaac 360
Db 460 TACCAATATGTGGACACCTTATACCAAGTGTGGAGGGTGTGGAGGACCGAGATGTGAACCCCAAC 519
Qy 361 cgtgagctgagcagagactgctgtacctcaacgtgtgagacacataccccgggctaca 420
Db 520 CBTGAGCTTGAGGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGGCTTACA 579
Qy 421 tcccccacccctgtctcgtctggtatcatatgggggtggtcttctacagtggggccctccc 480
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTTCTCC 639
Qy 481 ttggacgtatgaatgagcccttcttggtagagccgagagactgtgtgtccatg 540
Db 640 TTGGACGTGTACGATGGCGGCTTCTTGTGTACAGGCGGAGAGGACTGTGTGTGTCCATG 699
Qy 541 aactaccgggtggagccttggcttctcctgcttccctgcggggagccgagagccccgggc 600
Db 700 AACTACCGGGTGGGAGCCTTGTGGCTTCTGCGCCCTGCGGGGAGCCGAGAGCCCCGGGC 759
Qy 601 aatgtgggttctctctggaacagaggtggcctgcagtggtgaggaagaaagtgagcc 660
Db 760 AATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCGCTGAGTGGGTGCAGGAACTGGCAGCC 819
Qy 661 ttccgggggtgacccacatcagtagcagctgtttggggagagcggggagccgctcgtg 720
Db 820 TTCCGGGGGTGACCCACATCAGTACGCTGTTTGGGGAGAGCGCGGAGCCCTCGGTG 879
Qy 721 ggcacgacctgcttcccgccagccagcgggcctgttccacagggccgtgctgcagagc 780
Db 880 GGCATGCACTCTGCTCCCGCCCAAGCGGGCCCTGTTCACAGGGCCGTGTGCAGAGC 939
Qy 781 ggtgcccccaatggacccttggccacggtgggcatgggagagccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTTGGCCACGCTGGCATGGGATGGGAGGCCCGCTCGCAGGGCCACG 999
Qy 841 cagctggccccacctgtgggtctcctccagcgagcgacgtggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCACTGTGTGGGTGTCTCTCCAGCGGCACTGTGTGGGAATGACACAGAGCTG 1059
Qy 901 gttagctccttcgacacgacagcagagctcctggtgaacacgaatggcaactgtgtg 960
Db 1060 GTAGCTGCTCTTGGACACGACACGACGAGTCTGTGGTGAACACGAATGGCACGCTGTG 1119
Qy 961 cctcagaagaagcgtcttcgggtctcctcgtgctcgtggttagatggagacttctcagt 1020
Db 1120 CTTCAAGAAAGGCTTCCGGTTCCTCTGTCGCTGTGTGTAGATGGAGACTTCTCAGT 1179
Qy 1021 gacacccagagggccctcatcaacacgagagacttccacggcctcaggtgctgtgggt 1080
Db 1180 GACACCCAGAGGCCCTCATCAACGGGGAGACTTCCACGGGCTGCAGGTGTGTGTGGT 1239
Qy 1081 gtgggtgaaggaatgagggctgtatttctgtttacggggccccagggcttcaagcaagac 1140
Db 1240 GTGGTGAAGGAATGAGGGCTCGTATTTCTGTTTACGGGGCCCCAGGCTTACGAAGAAGAC 1299
Qy 1141 aacgagctctctcatcagcggcccgaggtctcctggcgggggtgcgggttcccccag 1200
Db 1300 AACGAGTCTCTCATCAGCGCGGCCAGAGTCTTCCGGGGGTGCGGGTTCGCCAG 1359

Qy 1201 gtaagtacccctggagccgagcggctgtgtctctgtaattacacagactggtgctgacccag 1260
Db 1360 GTAAGTACCTGGCAGCCGAGGCTGTGCTGTGCTGCAATTACACAGACTGGCTGCAATCCGAG 1419
Qy 1261 gaccggcacccctgagggagcggctgagcagatgtgtggcgacacacaaatgtctgtgc 1320
Db 1420 GACCCGGCAGCGCTGAGGAGGCCCTGAGCATGTGTGGGGGACCAATGTCTGTGTGC 1479
Qy 1321 cccgtggccccagctggctggggcagctggtgcccagggtgcccgggttctacgacctacgc 1380
Db 1480 CCGGTGGCCAGCTGGCTGGCGACTGTGCTGCCAGGCTGCCCGGTCTACGCCCTACGTC 1539
Qy 1381 ttgaaacaccgtgttccacgctctcctggccctgtgtgagtggggggtgcccccacggctac 1440
Db 1540 TTTGAACACCGTGTCTCCAGCTCTCTGGCCCTCTCTGGCCCTGTGATGGGGGTGCCCCACGGCTAC 1599
Qy 1441 gagatcgagttcatcttgggatccccctggacccctctcgaacctcacaacacagcagagag 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGCAGAGGAG 1659
Qy 1501 aaaatcttcgcccagcagctgactgactgactggtggccaaactttgcccgcacagggatccc 1560
Db 1660 AAAATCTTGGCCCGCAGGACTGATGCGATCTGGGCCAACTTTGCCCGCACACGGGATCCC 1719
Qy 1561 aatgagccccagagaccccccaaggcccccaatggtgcccccgctacacagcggggtcagcag 1620
Db 1720 AATGAGCCCCGAGACGCCCAAGGCCCAACAATGGCCCGCTACACGGCGGGGCTCAGCAG 1779
Qy 1621 tacgttagctgacactgcccgcgtgaggtgcgggggggtgcgggggggtgcggccccagcctgc 1680
Db 1780 TACGTTAGTGTGGAGCTGCGGGCCGTGGAGGTGCGGGGGGGGTGCGGGCCAGGCCCTGC 1839
Qy 1681 gctcttgaaacgcttctcctccccaaattgctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATGCTCAGGGCCACC 1881

RESULT 8

US-08-814-095-5
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/814.095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 3016 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Alternatively spliced Ache
ORIGINAL SOURCE: comprising exons 2, 3, 4, 5 and 6"
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 160..2010
US-08-814-095-5

Query Match 99.8%; Score 1722; DB 3; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccccagtgctgtctgacacagccttccctggttcccccactccttccctc 60
DB 160 ATGAGGCCCCCGCAGGTGCTGTGCACACGCTTCCTGGCTTCCCCACTCTCTCTC 219
QY 61 ctctctggtctcctggttgagagtggtgggctgagggccgggagagatgcagagctgctg 120
DB 220 CTCCTGTGGCTCCTGGGTGGAGAGTGGGGGCTGAGGGCGGGGAGATGCAGAGCTGCTG 279
QY 121 gtgacggtgctggtggcgctgctgagcattcgctaaagaccccggggcccctgtc 180
DB 280 GTGACGGTGGTGGGGCGCGCTGCGGGGCAATTCGGCTGAAGACCCCGGGGGCCCTGTG 339
QY 181 tctgcttctctgggcatcccttgcggagccaccatgggaccccgctgcgttcttgcca 240
DB 340 TCTGCTTTCTGTGGCATCCCTTTGCGGAGCCACCCATGGGACCCGCTGCTTCTGCCA 399
QY 241 ccggagcccaagcagccttggtcaggggtggtagacgttacaaccttcagagtgctgc 300
DB 400 CCGGAGCCCAAGCAGCTTGGTCAAGGGGTGTAGACGCTACAACTTCACAGAGTGTCTGC 459
QY 301 taccataatgtgagaccctataccacaggttttgaggccagcagatgtggaaccccaac 360
DB 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcagagactgctgtacctcaacgtgtgacacca tacccccgggctaca 420
DB 520 CGTGAAGCTGAGCGAGGACTGCTGTACCTCAACGCTGTGGACACCATACCCCGGCTACA 579
QY 421 tcccccaacctgtcctgtctgatactatgggggtggttctataagtggggacctctcc 480
DB 580 TCCCCCACCCCTGTCTCTGTGGATCTATGGGGGTGGCTTACAGTGGGGCTCCTCC 639
QY 481 ttgagcgtgtacagatggccgcttcttgggtacagccgagagactgtgctggtccatg 540
DB 640 TTGGAGCTGTACGATGGCGCTTCTTGTGTACAGGCCGACAGGACTGTGTGTGTCCATG 699
QY 541 aactaccgggtggagccttggcttctcctgcccctgcccggagccgagagccccgggc 600
DB 700 AACTACCGGTGGAGCCCTTGGCTTCTCTTGGCCCTCCTCCGGGAGCGGAGCGCCGGGC 759
QY 601 aatgtgggtctcctggtacagaggtgcccctgagtgagtggtgcagagaaagtggagcc 660
DB 760 AATGTGGGTCTCTGTGNTCAGAGGCTGGCCCTGCAGTGGGTGCAGGAAACGTGGCAGCC 819
QY 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcgcggagccctcgtg 720
DB 820 TTTCGGGGGTGACCCGACATCAGTGACCCCTGTTTGGGGAGAGCGCGGACCCGCTCGGTG 879
QY 721 ggcattgacctgtgttccccccagccgggctgttccacaggccctgctgcagagc 780
DB 880 GGCATGACCTGTGTGTCCCGCCAGCCGGGCTCTTCCACAGGSCCGTGTGTCAGAGC 939
QY 781 ggtgcccccaatggacacctggtggccacggttggtgcatgggagagcccgctgcagggccacg 840

DB 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGAGAGGCCCTGCCAGGGCCAG 999
QY 841 cagctggccccaccttgtgggctgtctcagggcgctcagtggtgggaatcacacagagcgtg 900
DB 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTCCAGGGCGGACACTGGTGGGAATCACACAGAGCTG 1059
QY 901 gtacgtgctcttgagacacagcagcgaggtcctggtgaacacacgaatggcactgtgctg 960
DB 1060 GTAGCCTGCTTCGGACACAGCACGACGAGGTCTGTGTGAACACACGATGGCAGCTGCTG 1119
QY 961 cctcaagaagaagcgtcttcgggttctcctctgtcctgtgctgtggtgagagacttccctcagt 1020
DB 1120 CCTCAAGAAGAGCGTCTTCGGGTTCCTTCGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
QY 1021 gacacccacagagccctcatcaacgcggagacttccacggcttcacggctcagtgctggtgggt 1080
DB 1180 GACACCCACAGAGCCCTCATCAACGGGGAGAGACTTCACGGGCTGCAGGTGCTGTGTGTGTGT 1239
QY 1081 gtggtgaaggaatgagggctcgtattttctggtttacggggggcccccaggtcttcagcaagac 1140
DB 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1299
QY 1141 aacgagctctcatcagccgggcccggaggttccctgcccggggtcggggttcccccag 1200
DB 1300 AACGAGTCTCTCATACAGCCGGCCGAGTTCTTGGCCGGGTGGGGTTCGGGTTCGCCAG 1359
QY 1201 gtaagtgaacctgagcagcagggctgtggtccttgcattacacagactggtgctgctcccgag 1260
DB 1360 GTAAGTGACCTGGCACCCGAGGCTGTGTCTTCATTTACACAGCTTGGCTGCATATCCCGAG 1419
QY 1261 gacccggcacgctcgtgagggagggccctgagcgtgtggtgggacacacaaatgctggtgctc 1320
DB 1420 GACCCGGCACGCTGTAGGGAGGGCCCTGAGCGATGTGTGGGGACACCAATGTCTGTGTGTC 1479
QY 1321 cccgtggccacgctggtcgtgggagactggtgcccaggtgcccgggtctacgacctacgtc 1380
DB 1480 CCGGTGGCCACGCTGTGCTGGGGACCTGGGTGCCAGGGTCCCGGGTCTACGCTTACGTC 1539
QY 1381 ttgaaacacgctgcttccacgctcctcctgcccctgtggtggtgggtgcccacaggtac 1440
DB 1540 TTGTGAACACCGTGTCTTCCAGCTCTCTTGGCCCTGTGTGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgtggttctccttgggatccccctggacccctctcgaacctcacacggcagagag 1500
DB 1600 GAGATCGAGTTCATCTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCGGGGCTCAGCAG 1659
QY 1501 aaaaatttcgcccagcagactgatactgggccaatttgcgcgcacagggggtatccc 1560
DB 1660 AAAATCTTCGCCACGCGACTGTATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagaccccccaagccccacaaatggcccccggtacacggcgggggtcagcag 1620
DB 1720 AATGAGCCCCGAGAGACCCCAAGGCCCCACAAATGGCCCCCTTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtcttgacactgcggcgcgtggtgaggtgctggggggtgctgcgccagggctgc 1680
DB 1780 TAGGTTAGTCTGGACCTGCGGCGCTGGAGGTGCGGCGGGGTGCGGCCCGCCAGGCTGC 1839
QY 1681 gcttctggaacgcgttctctcccaaatgtcctcagcgccacc 1722
DB 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATTTGCTCAGCGCCACC 1881

RESULT 9
US-08-318-826A-6
; Sequence 6, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense

;; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
;; TITLE OF INVENTION: Containing Them
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Kohn & Associates
;; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: US

;; ZIP: 48334

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/318,826A

;; FILING DATE:

;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Kohn, Kenneth I.

;; REGISTRATION NUMBER: 30,955

;; REFERENCE/DOCKET NUMBER: 2391.00001

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (248) 539-5050

;; TELEFAX: (248) 539-5055

;; INFORMATION FOR SEQ ID NO: 6:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 3096 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: double

;; TOPOLOGY: linear

;; MOLECULE TYPE: cDNA to mRNA

;; HYPOTHETICAL: NO

;; ANTI-SENSE: NO

;; ORIGINAL SOURCE:

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: 160..1959

;; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,

;; OTHER INFORMATION: 3, 4, 5 and the translated portion of Intron 4 (readthrough)"

US-08-318-826A-6

Query Match 99.88; Score 1722; DB 2; Length 3096;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggcccgagtgctgtctacacagcgttccctggttccctccctctctctc 60

Db 160 ATGAGGCCCGCAGTGTCTGTCTGCACACGCTTCCCTGGCTTCCCACTCTCTCTC 219

QY 61 ctctctggtctctgggtgagagagtggtggtgagggcggtgagagtgagagtgctg 120

Db 220 CTCCTCTGGCTCTCTGGGTGAGGAGTGGGGCTGAGGGCGGGAGGTGACAGACTGCTG 279

QY 121 gtgacgtgctggtggggcggtgctggggcattcgtcgaagacccccggggccctgtc 180

Db 280 GTGACGCTGCTGGGGCGGCTGCGGGCATTCGCTGAAGACCCCGGGGGCCCTGTC 339

QY 181 tctgcttctctggcatccctcttgcgagaccacccatggagaccctcgtcttctgcca 240

Db 340 TCTCTTCTCTGGGCATCTCCCTTTGCGGAGCCACCCATGGAGCCCGTCGCTTTCTGCCA 399

QY 241 ccgagagcccaagcagccttggttcaggggtggttagacgctacaaccttccagagtgctgc 300

Db 400 CCGAGAGCCCAAGCAGCTTGGTCAAGGGGTGTGACAGCTACAACTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtggacacccctatacccagggtttttgagggcaccgagatgtggaaaccccaac 360

Db 460 TACCAATATGTGGACACCTATACCCAGGTTTGTAGGGGACCGAGATGTGGAACCCCAAC 519

QY 361 cgtgagctgagcaggagactgctgtacctcaacgtgtgagacacataccccccggcctaca 420

Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGACACCATACCCCGGCTTACA 579

QY 421 tccccacccctctctctgctgtgatctatgggggtggtctctacagtggggcctccccc 480

Db 580 TCCCCACCCCTCTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTTCTTCC 639

QY 481 ttggacgtgtacgatggccgcttcttggtagacgccagagagactgtgtggttcccatg 540

Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGCGCCGAGAGACTGTGCTGCTGTGCTCATG 699

QY 541 aactacacgggtggagaccttggcttccctgcccctcccgagcggagccgagagccccgggc 600

Db 700 AACTACCGGGTGGAGACCTTTGGCTTCTGCTGCCCTGCCGGGAGCCGAGAGCCCCGGGC 759

QY 601 aatgtgggtctctggatcagaggtggccctcagtggtggtgaggaacagtgagcagcc 660

Db 760 AATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819

QY 661 ttccgggggtgaccggacatcagtgacgtgttgggggagagcgcggggagccgctcgtg 720

Db 820 TTCGGGGGTGACCCGACATCAGTACGCTGTTGGGGAGAGCGCGGGAGCCCTCGGTG 879

QY 721 ggaatgacactgctgtcccgcagccagccggggcctgttccacagggccgctgctcagagc 780

Db 880 GGCATGCACTGCTGTCCCCGCCAGCGGGGCTGTGTTCCACAGGGCGCTGCTGCAGAGC 939

QY 781 ggtgcccccaatgacctggccaggtggcagtggtggcagtgagagggccgctcagagccacg 840

Db 940 GGTGCCCCCAATGAGACCTGGGCCACGCTGGGATGGAGAGGGCCCTGCGAGGGCCACG 999

QY 841 cagctggcccaaccttgggtgctctccagcggcagctggtgggaatgacacagagctg 900

Db 1000 CAGCTGGCCCACTTGTGGGTGCTCTCCAGCGGCACACTGTGTGGGAATGACACAGCTG 1059

QY 901 gtacgtcctctcggacacagcagcagcagcagcagcagcagcagcagcagcagcagcag 960

Db 1060 GTAGCTGCTCTCGGACACGACACGACGCTCTGTGTGTAACACCAAGATGACAGCTGTG 1119

QY 961 cctcaagaaagcctctctcgttctctcgtcgtcgtggtggtggtggtggtggtggtggt 1020

Db 1120 CCTCAAGAAAGCTTCTCCGGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179

QY 1021 gacccccagagggccctcctcaacgcgggagacttccacggcctcgagggctgctggtggtg 1080

Db 1180 GACACCCAGAGGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGTGTGGT 1239

QY 1081 gtggtgaagatgagggctcgtatttctgttttacggggcccccagcgttcagcaaaagac 1140

Db 1240 GTGTTGAAGGATGAGGGCTCTGTTTCTGTTTACGGGGCCCCAGGCTTTCAGCAAGAC 1299

QY 1141 aacgagctctctcagcggggcgagttcctggccgggtggtggtggtggtggtggtggtggt 1200

Db 1300 AACGAGTCTCTATCAGCCGGCGCGAGTTCCTGGCCGGGGTGGGGGTGCGGGTTCCTCCAG 1359

QY 1201 gtaagtaccttggcagccgaggtggtggtcgtcattacacagactggtcgtcgtcgtcgtc 1260

Db 1360 GTAAGTACCTTGGCAGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1419

QY 1261 gaccggcagccctgagggagggcctgagcagatgtgtgtggcgagaccacaaatgtcgtgtgc 1320

Db 1420 GACCGGCACCCCTGAGGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1479

QY 1321 cccgtggccagcgtggtggcgagcgtggtggtggtggtggtggtggtggtggtggtggtggt 1380

Db 1480 CCCGTGGCCAGCTGCTGGTGGCGGACTGGCTGCCAGGGTGGCCGGGTGTACGCCCTACGTC 1539

QY 1381 ttgaaacacgtgttccacagcctcctcgtggtggtggtggtggtggtggtggtggtggtggt 1440

Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTCTGCGCCCTGTGGATGGGGTGTGCCCCAGCGCTAC 1599

QY 1441 gagatcaggttcattcttttgggatccccctggacccctctcgaacactcagacagagagag 1500

Db	1600	GAGATCGAGTTCACTTTTGGGATCCCCCTGGACCCCTCTCAAACTACACGGCAGAGGAG	1659
Qy	1501	aaaattcttcgcccagcagctgtagtcgatactatggtggccaactttgccccgcacaggggatcccc	1560
Db	1660	AAAATCTTTGGCCCGAGGACTGATGCGATACTGGGCCAACTTTGGCCCGCAGACGGGATCCC	1719
Qy	1561	aatgagccccgcgagaccccaaggcccccaaatggcccccgctacacggtcggggggtctcaagcag	1620
Db	1720	AATGAGCCCCCGAGAGCCCAAGGCCCCCAATGGCCCCGTACACGCGGGGGCTCAGCAG	1779
Qy	1621	tacgttagcttggaacctgcggtcgctgtaggtgtagggggggtctgcgcgccaggcctgc	1680
Db	1780	TAGTTTAGTCTGAGACTGCGGCGCGCTGAGAGGTGCGGCGGGGGGCTGCGGCGCCAGGCGCTGC	1839
Qy	1681	gcattctgaaacgcgttctctcccaaaattgctcagcgcacc	1722
Db	1840	GCCTCTTGAAACCGCTTCTCTCCCAAAATTTGCTCAGCGCGCAC	1881

RESULT 10

```

US-08-370-156-3
:
: Sequence 3, Application US/08370156
: Patent No. 5932780
: GENERAL INFORMATION:
: APPLICANT: Soreq, Hermona
: APPLICANT: Zakut, Haim
: APPLICANT: Shani, Moshe
: TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
: TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Reising, Ethington, Barnard & Perry
: STREET: P.O. Box 4390
: CITY: Troy
: STATE: Michigan
: COUNTRY: US
: ZIP: 48099
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/370,156
: FILING DATE:
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Kohn, Kenneth I.
: REGISTRATION NUMBER: 30,955
: REFERENCE/DOCKET NUMBER: P-307 (Mulford)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (810) 689-3500
: TELEFAX: (810) 689-4071
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3096 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 160..1959
:
US-08-370-156-3

```

```
Query Match          99.8%; Score 1722; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 atgagggcccccgcagtgctgctgcacacgccttcctggtccctcccaactcctctctc 60
|||||
Db 160 ATGAGGGCCCCCGCAGTGCTGCTGCACACGGCTTCCTGGCTGCCCATCCCTTCCTC 219

Db 880 GGATGACCTGCTGTCGCCGCCACCGGGGCTTTCACAGGGCCCTGCTGCAGAC 939
QY 781 ggtgcccccaatgacctggtggccacagtggtgcatgagagcccgctgcagggccag 840
Db 940 GGTGCCCCCAATGAGACCTTGGCCACCGTGGGATGGAGAGCCCGCTGCAGGGCCACG 999
QY 841 cagctggccacactgtggtgctctccagcggcactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCACCTTGTGGCTGTCCTCCAGCGGCACATTCGGTGAATGACACAGAGCTG 1059
QY 901 gtacgtgctcttgagacagacagcagcaggtctgtgtaacacacaaatggacgtgctg 960
Db 1060 GTAGCTGCTTGGGACAGACAGCGCAGGTCTCTGTTGAACACCAATGGCACGTGCTG 1119
QY 961 cctcaagaagcgtctccggtctctctgctgctgtgtagtagagactctccagt 1020
Db 1120 CCTCAAGAAGCGTCTCCGGTCTCTCTGCTGCTGTGTGTAGTAGGACATTCCTCAGT 1179
QY 1021 gacacccagagccctcatcaacgcggagacttccacggcctgcagtgctggtgggt 1080
Db 1180 GACACCCAGAGCCCTCATCAACGGGGGAGACTTCCACGGCTGCAGTGTGTGGGT 1239
QY 1081 gtggtgaagatgagggctcgattttctggtttacagggggcccaaggtctcagcaaaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTCTGTGTTACGGGGGCCCGACGGCTTCAGCAAGAC 1299
QY 1141 aacgagctctcatcagcggcggcggagtcctgcccgggtgcgggtggttcccag 1200
Db 1300 AAGGAGTCTCATACAGCGGGCCGAGTCTCTGGCCGGGGTGGGGTTCGCCAG 1359
QY 1201 gtaagtacctggcagccaggtgctgtgcttcgtcattacacagactggtgcatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGTCTCTGCAATACACAGACTGGCTGCATCCGAG 1419
QY 1261 gaccggcagcctgagggagccctgagcagtggtgtggggagacacaaatgctgtgac 1320
Db 1420 GACCCGGCAGCCTGAGGGAGGCCCCGTGAGCGATGTGGTGGGGGACCAATGTCGTGTGC 1479
QY 1321 ccgtgcccagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1380
Db 1480 CCGTGGCCAGCTGCTGGCGACCTGGCTGCCAGGGTGGCGGGTCTACGCCCTACGTC 1539
QY 1381 ttgaaacacgtgcttccacgtctctctgcccctggtggtggtggtggtggtggtggtggt 1440
Db 1540 TTTGAACACCGCTGTTCCACGCTCTCTGGCCCTCTGATGGGGTGGCCACCGCTAC 1599
QY 1441 gagatgagttcatcttggatccccctggaccctctcgaactacacgacagagag 1500
Db 1600 GAGATGAGTTTATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGGAG 1659
QY 1501 aaaaattctgcgcagcactgactgatactggtgcaactttgcccgcacaggggatccc 1560
Db 1660 AAAATCTTGGCCAGGACTGATGCGATACTGGGCCAATTTGCCCGCACGGGATCCC 1719
QY 1561 aatgagcccgagaccacaaagccccacaaatggccccctacacggcgggctcagcag 1620
Db 1720 AATGACCCCGAGACCCCAAGGCCCCACAAATGGCCCGCTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagttcagacctgcgcccgtgaggtgagcggggggtgcgcccagggcctgac 1680
Db 1780 TAGGTTAGTCTGAGACCTGCGCCGCTGGAGGTGCGCGGGGGTGGCGCCCGAGGCTGC 1839
QY 1681 gctcttggaaaccgtctctcccaaatgctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCCGCCAAATGCTCAGCGCCACC 1881

RESULT 12

US-08-814-095-7

; Sequence 7, Application US/08814095

; Patent No. 6025183

; GENERAL INFORMATION:

APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Shani, Moshe
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 No. 6025183thwestern Highway, Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 35060 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Cosmid including ACHE
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 7q22
FEATURE:
NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHE Promotor"
OTHER INFORMATION: /standard_name= "ACHE Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon

QY 481 ttgagactgtacagatggccgctctcttggtacagggccagagagactgtgctggttccatg 540
Db 24590 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGCCGACAGGACTGTGCTGGTCCATG 24649
QY 541 aactacgggtggagaccttggcttctccctggccctccgggagccagagcccccggc 600
Db 24650 AACTACGGGTGGAGCCCTTGGCTTCCCTGGCCCTCCGGGGAGCCGAGAGCCCGCCGGC 24709
QY 601 aatgtgggtctctctggtatcagagggctggccctgagtggtgaggaacagtggtcagcc 660
Db 24710 AATGTGGGTCTCTCTGATCAGAGGCTGGCCCTGCACTGGGTGAGGAACTGGCAGCC 24769
QY 661 ttggggggtagcccgacacatcagtgacgctgtttggggagagccgagggccctcggtg 720
Db 24770 TTGGGGGGTAGCCCGACATCAGTGACGCTGTGTTGGGGAGAGCGCGGCGCTCGGTG 24829
QY 721 ggcatacactgctgtcccccgcacagccgggctcttccacagggcctgctgcagagc 780
Db 24830 GGCATGCACTGTGTCTCCCGCCCGCCGCGGGGCTGTTCACAGGGCGCTGCTGCAGGC 24889
QY 781 ggtgcccccaatggagccctgggcccaggtgggcatgggagagcccgctgcagggccacg 840
Db 24890 GGTGCCCCCAATGGACCTTGGGCCACGGTGGGCATGGAGAGGCCGTCGACGGCCACG 24949
QY 841 cagctggccacactgtggtgctctctccagggcagctggtgggaatgacacagagctg 900
Db 24950 CAGCTGGCCACCTTGTGGGCTGTCTCCAGGGCGCACTGGTGGGAATGACACAGAGCTG 25009
QY 901 gtacgtgctctggagacagccagcagcagctgctgtgtaaacacagagcagctgctg 960
Db 25010 GTACCTGCTCTCGACACGACGAGCGGAGGCTGCTGTGAACCCAGAAATGGACAGCTG 25069
QY 961 cctcaagaagcgtcttccgggttctctctgctgctgctgtagatggagacttcctcagt 1020
Db 25070 CCTCAAGAAGCGCTTCCGGTGTCTCTCTGCTGCTGTGTAGATGGAGACTTCTCTCAGT 25129
QY 1021 gacacccagagccctcatcaacgagggagagacttcacagggcctgcaggt 1070
Db 25130 GACACCCAGAGCCCTCATCAACGGGGAGAGCTTCCACGGCGCTGCAGGT 25179

RESULT 13

US-08-370-156-13
; Sequence 13, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethlington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-156-13

Query Match 2.0%; Score 35; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1363 cgggtctagcctcagctcttcttgacacccgtgcttc 1397
Db 1 CGGGTCTACGCTACGCTCTTTGACACCCGTGCTTC 35

RESULT 14

US-08-370-156-23/c
; Sequence 23, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethlington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-156-23

Query Match 2.0%; Score 35; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1604 cggcgggggctcagcagtagcttagtctgacctg 1638
Db 35 CGCGGGGGGCTCAGCAGTAGCTAGTCTGACCTG 1

RESULT 15

```

US-08-875-710-4
; Sequence 4, Application US/08875710
; Patent No. 6326139
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; TITLE OF INVENTION: METHOD OF SCREENING FOR GENETIC PREDISPOSITION TO
; TITLE OF INVENTION: ANTICHOLINESTERASE THERAPY
; FILE REFERENCE: 2391.00076
; CURRENT APPLICATION NUMBER: US/08/875,710
; CURRENT FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: PCT/US96/00322
; EARLIER FILING DATE: 1996-01-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-875-710-4

```

```

Query Match      2.08; Score 35; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1363  cgggtctacgcctacgtctttgaacacccgtgcttc 1397
          |||||
Db 1      cgggtctacgcctacgtctttgaacacccgtgcttc 35
          |||||

```

Search completed: August 31, 2002, 22:45:22
Job time: 19786 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:51:31 ; Search time 824.51 Seconds
(without alignments)
12008.885 Million cell updates/sec

Title: US-09-810-861b-3
Perfect score: 5767
Sequence: 1 agctgcagctcgagtc.....ctatgaccatgattacgcca 5767

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5767	100.0	5767	24	Plasmid pTM034. S
2	2637	45.7	3858	22	plasmid PRK50 used
3	2637	45.7	4960	22	PCMV-I-Cre-pA vect
4	2637	45.7	7332	22	PCMV-I-beta-pA vec
C 5	2637	45.7	8062	22	Plasmid PRK76 used
C 6	2637	45.7	8153	22	Plasmid PRK74 used
7	2637	45.7	8854	20	DNA encoding chimera
C 8	2630	45.6	6295	19	Plasmid pWRG3196 e
C 9	2630	45.6	7164	22	Commercial plasmid

C 10	2630	45.6	7383	22	AAS00153
C 11	2628	45.6	4950	18	AAV03801
C 12	2628	45.6	4950	22	AAV03801
C 13	2628	45.6	5733	21	AAA95416
C 14	2628	45.6	5733	21	AAA95416
C 15	2628	45.6	9632	20	AAV04251
C 16	2628	45.6	9632	20	AAV04251
C 17	2628	45.6	9632	20	AAV04251
C 18	2628	45.6	11846	20	AAV04251
C 19	2586	44.8	12022	22	AAV04251
C 20	2577	44.7	4163	13	AAQ32349
C 21	2534	43.9	5897	20	AAV63741
C 22	2534	43.9	5897	20	AAV63741
C 23	2534	43.9	5897	20	AAV63741
C 24	2533	43.9	3753	24	ABA04129
C 25	2533	43.9	3753	24	ABA04129
C 26	2533	43.9	4283	18	AAV04130
C 27	2533	43.9	4283	18	AAV04130
C 28	2533	43.9	4283	19	AAV04130
C 29	2533	43.9	4283	21	AAV04130
C 30	2533	43.9	6971	22	AAV55126
C 31	2533	43.9	7558	22	AAV55126
C 32	2533	43.9	7969	22	AAV55126
C 33	2528	43.8	4776	20	AAV77617
C 34	2528	43.8	4776	20	AAV77617
C 35	2527	43.8	4776	20	AAV77617
C 36	2527	43.8	7639	14	AAQ42160
C 37	2526	43.8	4713	19	AAQ42160
C 38	2526	43.8	4713	19	AAV12067
C 39	2526	43.8	4724	19	AAV12068
C 40	2526	43.8	5421	21	AAV68299
C 41	2526	43.8	5421	21	AAV68299
C 42	2526	43.8	6116	21	AAV68299
C 43	2526	43.8	6708	21	AAV68299
C 44	2509	43.5	7380	20	AAV68299
C 45	2480	43.0	4045	15	AAQ70942

ALIGNMENTS

RESULT 1
AAS17547
ID AAS17547 standard; DNA; 5767 BP.
AC AAS17547;
DT 25-FEB-2002 (first entry)
DE Plasmid pTM034.
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW circular; ds.
XX Synthetic.
XX OS
XX WO200171014-A2.
XX 27-SEP-2001.
XX 16-MAR-2001; 2001WO-US08468.
XX 17-MAR-2000; 2000US-190440P.
XX (MORT/) MOR T.
XX (SORE/) SOREQ H.
XX (ARNT/) ARNTZEN C.
XX (MASO/) MASON H.
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

Mor T, Soreq H, Arntzen C, Mason H:

WPI: 2002-055120/07.

Production of a transgenic plant which contains a polynucleotide that encodes a human acetylcholinesterase which upon purification is effective against acetylcholinesterase poisoning -

Claim 11: page 28-31; 42pp: English.

The invention relates to a method of producing a transgenic plant which contains a polynucleotide that encodes human acetylcholinesterase (AChE) which upon purification is effective against acetylcholinesterase poisoning. The method is used for treating a victim of acetylcholinesterase poisoning by administering a therapeutic amount of a physiologically active human acetylcholinesterase expressed in plant tissue. The extensive use of anticholinesterase pesticides with concurrent accidental poisoning, the threat of chemical warfare and environmental concerns demand the development of effective, inexpensive and stage countermeasures and bioremediation solutions. Prior art methods for treating AChE poisoning have used the muscarinic receptor antagonist atropine and oximes to reactivate the organophosphate(OP)-modified AChE. The reversible carbamate, pyridostigmine bromide has also been used as a prophylactic. However, these conventional treatments have limited effectiveness and serious short and long-term side effects and may result in significant performance deficits and even permanent brain damage. This invention permits the utilisation of cholinesterases to counter-act the toxic effects of anti-cholinergic agents, using transgenic plants for the production of the enzymes is cost effective and the product is stable so that the injected enzymes have the advantage of having a long half-life. The transgenic form of the enzymes are also easy to purify. The present sequence is plasmid pPM034, the pGPTKan derivative construct used in the generation of transgenic tomato plants that constitutively express human AChE.

Sequence 5767 BP: 1390 A: 1521 C: 1495 G: 1361 T: 0 other:

Query Match	100.0%	Score 5767:	DB 24:	Length 5767:
-------------	--------	-------------	--------	--------------

Query Match	100.0%;	score 3767;
Best Local Similarity	100.0%;	Pred. No. 0;

Best local similarity 100.0%; PIED. NO. 0;
Matches 5767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	agttgcattgcctgcagggtcaacatggtggagcacgacactctcgtctactccaagaata	60
Db	1	agttgcattgcctgcagggtcaacatggtggagcacgacactctcgtctactccaagaata	60
QY	61	tcaaaatcacgctctcagaagaccagagggctattgagacttttcacaaaagggttaatat	120
Db	61	tcaaaatcacgctctcagaagaccagagggctattgagacttttcacaaaagggttaatat	120
QY	121	cgggaaacctcttcggattccattgcccagctatctgtcacttcatcgaagagcagtag	180
Db	121	cgggaaacctcttcggattccattgcccagctatctgtcacttcatcgaagagcagtag	180
QY	181	aaaggaagaatggcttctcaaatgccatcattgcgataaagaaaggctatcgtttcaag	240
Db	181	aaaggaagaatggcttctcaaatgccatcattgcgataaagaaaggctatcgtttcaag	240
QY	241	aatgcctctaccgacagatgggtcccaagaatgacccccccacgacgaggaacatcgtggaa	300
Db	241	aatgcctctaccgacagatgggtcccaagaatgacccccccacgacgaggaacatcgtggaa	300
QY	301	aaagagaagctttccaaaccgctcttccaaagcaagtggattgattgataaacttttcaaca	360
Db	301	aaagagaagctttccaaaccgctcttccaaagcaagtggattgattgataaacttttcaaca	360
QY	361	aagggtaatatcgggaaacctctcgattccattgccagctatctgtcacttcatcga	420
Db	361	aagggtaatatcgggaaacctctcgattccattgccagctatctgtcacttcatcga	420
QY	421	aagggacagtagaaaggaagatgggtctctcaaaatgccatctattgcgataaagggaagggc	480
Db	421	aagggacagtagaaaggaagatgggtctctcaaaatgccatctattgcgataaagggaagggc	480

Qy 1561 atgcacotgctgtcccccagccagccggcctgtttccacagagggcgctgctcagagcgggt 1620
Db 1561 atgcacotgctgtcccccagccagccggcctgtttccacagagggcgctgctcagagcgggt 1620
Qy 1621 gcccccgaatggacccctgggcccagcgggtggatggagagggcccgctcgagagggccacgcag 1680
Db 1621 gcccccgaatggacccctgggcccagcgggtggatggagagggcccgctcgagagggccacgcag 1680
Qy 1681 ctggcccacacttgctgggctgtcctccagggcgccactgggtgggaatgacacagagctggta 1740
Db 1681 ctggcccacacttgctgggctgtcctccagggcgccactgggtgggaatgacacagagctggta 1740
Qy 1741 gccctgtctggacacacacacagccaggtcctgtgtaaccacgaatggcacgtctgcct 1800
Db 1741 gccctgtctggacacacacacagccaggtcctgtgtaaccacgaatggcacgtctgcct 1800
Qy 1801 caagaaagcgtctcccggttctccttctcgtcctgtgtgtagatggagacttccctcagtgac 1860
Db 1801 caagaaagcgtctcccggttctccttctcgtcctgtgtgtagatggagacttccctcagtgac 1860
Qy 1861 accccagagggccctcaataacacggggagacttccacggcctgcagggctgctgtgggtgtg 1920
Db 1861 accccagagggccctcaataacacggggagacttccacggcctgcagggctgctgtgggtgtg 1920
Qy 1921 gtgaagatgaggtcgtctatttctgttttacggggcccgagcttcagcaaaagacaac 1980
Db 1921 gtgaagatgaggtcgtctatttctgttttacggggcccgagcttcagcaaaagacaac 1980
Qy 1981 gagtctctcaacagccggcgaggttctcgtccgggggtgcgggttcccccagta 2040
Db 1981 gagtctctcaacagccggcgaggttctcgtccgggggtgcgggttcccccagta 2040
Qy 2041 agtgacctggcagccgagggctgtgtcctgcattacacagactggctgcatcccgaggac 2100
Db 2041 agtgacctggcagccgagggctgtgtcctgcattacacagactggctgcatcccgaggac 2100
Qy 2101 ccgacagcctgagggagggccctgagcagatggtggcgacacaatgtcgtgtgccc 2160
Db 2101 ccgacagcctgagggagggccctgagcagatggtggcgacacaatgtcgtgtgccc 2160
Qy 2161 gtggcccagctgggtggcgactggctgcccaggggtgcccgggtctacgctcagctcttt 2220
Db 2161 gtggcccagctgggtggcgactggctgcccaggggtgcccgggtctacgctcagctcttt 2220
Qy 2221 gaacacgtgtttccacgctctcgtgcccctgtggatgggggtgcccacggctacag 2280
Db 2221 gaacacgtgtttccacgctctcgtgcccctgtggatgggggtgcccacggctacag 2280
Qy 2281 atcgagttcatctttgggattcccctgagccctctcgaaactacacggcagagagaaa 2340
Db 2281 atcgagttcatctttgggattcccctgagccctctcgaaactacacggcagagagaaa 2340
Qy 2341 atcttcccagcactgatgcatactgggcccacttggcccacaggggattcccact 2400
Db 2341 atcttcccagcactgatgcatactgggcccacttggcccacaggggattcccact 2400
Qy 2401 gagcccagagaccacagggcccacaaatggcccctgtacacggcggggctcagcagtac 2460
Db 2401 gagcccagagaccacagggcccacaaatggcccctgtacacggcggggctcagcagtac 2460
Qy 2461 gttagttcgtgacctggcgccgtgaggtgcgcgggggctgcgcgccagggcctgcgc 2520
Db 2461 gttagttcgtgacctggcgccgtgaggtgcgcgggggctgcgcgccagggcctgcgc 2520
Qy 2521 ttctggaaacctctctcccacaaatgtctcagcgtactcgtataggtaccgagctctctc 2580
Db 2521 ttctggaaacctctctcccacaaatgtctcagcgtactcgtataggtaccgagctctctc 2580
Qy 2581 aacaaatagtagtattgtctcctataatataatgaaaggtatgctgatatgactat 2640
Db 2581 aacaaatagtagtattgtctcctataatataatgaaaggtatgctgatatgactat 2640
Qy 2641 tcaaataggagcattagctatgtttgttaatgtcaccttattgttatgtgggtaagtcacc 2700

Db 2641 tcaaataggagcattagctatgtttgttaatgtcacttattgttatgtgggtaagtcacc 2700
Qy 2701 taagacactcacgtacactacgttgttctcttaccggctttaaataactctctgcct 2760
Db 2701 taagacactcacgtacactacgttgttctcttaccggctttaaataactctctgcct 2760
Qy 2761 tgttccattattactaattatccctttctcactaaagaaaattgttatcattaaagtat 2820
Db 2761 tgttccattattactaattatccctttctcactaaagaaaattgttatcattaaagtat 2820
Qy 2821 tagcttttagaacatatagagcttttaattgggtagggtttttacaaattaaactaaataaa 2880
Db 2821 tagcttttagaacatatagagcttttaattgggtagggtttttacaaattaaactaaataaa 2880
Qy 2881 atgtcataaaatccacgtggtttaacaaaatgcagaaaaatcgactcgtctatttgaccga 2940
Db 2881 atgtcataaaatccacgtggtttaacaaaatgcagaaaaatcgactcgtctatttgaccga 2940
Qy 2941 cagttgtctatttaataatggggccaccatagtagactgacaaaataaattacctgacaca 3000
Db 2941 cagttgtctatttaataatggggccaccatagtagactgacaaaataaattacctgacaca 3000
Qy 3001 tccgtttcactaaaataacacacacaaaaggagtcattttccaggcgcattttttaata 3060
Db 3001 tccgtttcactaaaataacacacacaaaaggagtcattttccaggcgcattttttaata 3060
Qy 3061 aaaaacagttaaaaggagtgcaatagaaaatagggtgtggaaatagttgattgagca 3120
Db 3061 aaaaacagttaaaaggagtgcaatagaaaatagggtgtggaaatagttgattgagca 3120
Qy 3121 cgtcttgaaagcgaattcactggcctgtttttacaacgtcgtgactgggaaacccctggc 3180
Db 3121 cgtcttgaaagcgaattcactggcctgtttttacaacgtcgtgactgggaaacccctggc 3180
Qy 3181 gttacccaaactaaatcgcttgcagcacatccccctttccgcagctggcgtaataagcaa 3240
Db 3181 gttacccaaactaaatcgcttgcagcacatccccctttccgcagctggcgtaataagcaa 3240
Qy 3241 gagggccgcacacgatcgcccttcccaacagttggcgacgctgaaatggcgccctg 3300
Db 3241 gagggccgcacacgatcgcccttcccaacagttggcgacgctgaaatggcgccctg 3300
Qy 3301 atgcggtattttctccttacgcatctgtgcggtattttcacacggcattatggtgcactctc 3360
Db 3301 atgcggtattttctccttacgcatctgtgcggtattttcacacggcattatggtgcactctc 3360
Qy 3361 agtacaattctcgtctgacgcgcatagttaaagccagcccgacaccccgccacacccgct 3420
Db 3361 agtacaattctcgtctgacgcgcatagttaaagccagcccgacaccccgccacacccgct 3420
Qy 3421 gacggccctgacgggctgtctgctcccgcatcccgcttacagacaagctgtgaccgtc 3480
Db 3421 gacggccctgacgggctgtctgctcccgcatcccgcttacagacaagctgtgaccgtc 3480
Qy 3481 tccgggagctgcattgtcgaagggttttccacgctcactacacgaaacgcgcgagacgaag 3540
Db 3481 tccgggagctgcattgtcgaagggttttccacgctcactacacgaaacgcgcgagacgaag 3540
Qy 3541 ggcctcgtgacgcctatttttattaggttaattgtcagataaataatggtttcttagacg 3600
Db 3541 ggcctcgtgacgcctatttttattaggttaattgtcagataaataatggtttcttagacg 3600
Qy 3601 tcagggtgcaacttttcggggaaaatgtgcgcggaacccctattttttatttttctaaata 3660
Db 3601 tcagggtgcaacttttcggggaaaatgtgcgcggaacccctattttttatttttctaaata 3660
Qy 3661 cattcaataatgtatccgctcatgagacaataaacctgataaaatgcttcaataattga 3720
Db 3661 cattcaataatgtatccgctcatgagacaataaacctgataaaatgcttcaataattga 3720
Qy 3721 aaaagagagatagagttacaaacatttccggtgtcgcgccttatttcccttttttgcgca 3780

Db	3721	aaaggaagatgatgatattcaacatttcgctgcgcccttatcccttttttgcggca	3780
Qy	3781	tttgcctctctgtttttgtctaccagaaaacgctgggtgaagtaaaagatgctgaagat	3840
Db	3781	tttgcctctctgtttttgtctaccagaaaacgctgggtgaagtaaaagatgctgaagat	3840
Qy	3841	cagttgggtgcacgagtggtttacatcgaaatgdatctcaacacgctgaagatcccttgag	3900
Db	3841	cagttgggtgcacgagtggtttacatcgaaatgdatctcaacacgctgaagatcccttgag	3900
Qy	3901	agttttgcgcccgagaagacgttttccaatgatgagcaacttttaaaagtctgctatgtggc	3960
Db	3901	agttttgcgcccgagaagacgttttccaatgatgagcaacttttaaaagtctgctatgtggc	3960
Qy	3961	gcggtattatcccgattgacgcgggcaagagcaactcggctgcccaatacactattct	4020
Db	3961	gcggtattatcccgattgacgcgggcaagagcaactcggctgcccaatacactattct	4020
Qy	4021	cagaaatgaacttgggttgactacacacgctccagaaaaagcatcttaacggtggcatgaca	4080
Db	4021	cagaaatgaacttgggttgactacacacgctccagaaaaagcatcttaacggtggcatgaca	4080
Qy	4081	gtaagagaattatgcagtgctgcctaaccataaccatgagtgataacactgcggccaacttaatt	4140
Db	4081	gtaagagaattatgcagtgctgcctaaccataaccatgagtgataacactgcggccaacttaatt	4140
Qy	4141	ctgacaacgatcggaggaacccgaagagcatcaccgcttttttgacaacataggggatcat	4200
Db	4141	ctgacaacgatcggaggaacccgaagagcatcaccgcttttttgacaacataggggatcat	4200
Qy	4201	gtaactcgccttgatcgtttgggaacccgagctggaatgaagccaaccacaacgacgagcgt	4260
Db	4201	gtaactcgccttgatcgtttgggaacccgagctggaatgaagccaaccacaacgacgagcgt	4260
Qy	4261	gacacacacgatgcctgtagcaatggcaacaacgtttgcgcaaacatttaactggcgaaacta	4320
Db	4261	gacacacacgatgcctgtagcaatggcaacaacgtttgcgcaaacatttaactggcgaaacta	4320
Qy	4321	cttactctagcttcccgcgaacaaataatagactggatggagcgcgataaagtgtgaagga	4380
Db	4321	cttactctagcttcccgcgaacaaataatagactggatggagcgcgataaagtgtgaagga	4380
Qy	4381	ccacttctgcgctcggcccttcggctggctgggtttattgtcgtgataaactctggagccggt	4440
Db	4381	ccacttctgcgctcggcccttcggctggctgggtttattgtcgtgataaactctggagccggt	4440
Qy	4441	gagcgtgggtctgcggtatcatctgacacactggggccagatgtaagccctcccgatc	4500
Db	4441	gagcgtgggtctgcggtatcatctgacacactggggccagatgtaagccctcccgatc	4500
Qy	4501	gtagttatctacacacggggagtcaggcaactatggatgaacgaaatagacagatcgct	4560
Db	4501	gtagttatctacacacggggagtcaggcaactatggatgaacgaaatagacagatcgct	4560
Qy	4561	gagatagggtgcctcactgatgaagcatgtgaactgtcagaccgaagttaactcatatata	4620
Db	4561	gagatagggtgcctcactgatgaagcatgtgaactgtcagaccgaagttaactcatatata	4620
Qy	4621	ctttagattgattaaaaactcattttkatttaaaaggatctagggtgaagatcccttttt	4680
Db	4621	ctttagattgattaaaaactcattttkatttaaaaggatctagggtgaagatcccttttt	4680
Qy	4681	gataactcatgacaaaaatcccttaacgctgagtttttcgcttccactgagcgctcagacccc	4740
Db	4681	gataactcatgacaaaaatcccttaacgctgagtttttcgcttccactgagcgctcagacccc	4740
Qy	4741	gtagaaaaagatcaaaaggaacttctcttgagatcctttttctgcgcgtkactctgctctg	4800
Db	4741	gtagaaaaagatcaaaaggaacttctcttgagatcctttttctgcgcgtkactctgctctg	4800
Qy	4801	caaaacaaaaaaccccgctacacgagtggtttgtttgtccgggatacgaagctaccacact	4860
Db	4801	caaaacaaaaaaccccgctacacgagtggtttgtttgtttgtccgggatacgaagctaccacact	4860

RESULT	2
AAD04947	
ID	AAD04947 standard; DNA; 3858 BP.
XX	
AC	AAD04947;
XX	
DT	17-JUL-2001 (first entry)

XX Plasmid pRK50 used to test Cre recombinase mediated inversion.
 XX Gene trapping construct; conditional mutation; unidirectional inversion;
 KW recombinase recognition sequence; RRS; disruption cassette;
 KW selection cassette; transgenic organism; pRK50; Cre recombinase; ds.
 XX Chimeric - Adenovirus.
 OS Chimeric - Unidentified.
 OS Chimeric - ECMV virus.
 XX WO200129208-A1.
 XX 26-APR-2001.
 XX 16-OCT-2000; 2000WO-EP10162.
 XX 16-OCT-1999; 99EP-0120592.
 PR 27-OCT-1999; 99US-0162016.
 XX (ARTE-) ARTEMIS PHARM GMBH.
 PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
 XX Kuehn R, Von Melchener H, Altschmied J;
 PI WPI; 2001-308486/32.
 XX New gene trapping construct capable of causing conditional mutations in
 PT genes, comprises functional DNA segment inserted in sense or antisense
 PT direction relative to gene to be trapped -
 XX Example 3; Page 73-74; 78pp; English.
 XX The present invention relates to a conditional gene trapping construct
 CC capable of causing conditional mutations in genes. The gene trapping
 CC construct comprises two functional DNA segments, each being flanked by
 CC two recombinase recognition sequences (RRSs) specific to site specific
 CC recombinase which is capable of unidirectional inversion of double
 CC standard DNA segment. One of the DNA segment (disruption cassette) is
 CC inserted in antisense orientation relative to the transcriptional
 CC orientation of the gene to be trapped. The other DNA segment (selection
 CC cassette) is inserted in sense direction relative to the transcriptional
 CC orientation of the gene to be trapped. The cell comprising the gene
 CC trapping construct is useful for the identification and/or isolation of
 CC genes. The transgenic organism comprising the gene trapping construct is
 CC useful to study gene function at various developmental stages. The gene
 CC trapping construct is useful for mutationally inactivating all cellular
 CC genes. The present sequence is pRK50 vector, which is used to test Cre
 CC recombinase mediated inversion, which is related to the invention.
 XX Sequence 3858 BP; 943 A; 959 C; 994 G; 962 T; 0 other;

Query Match 45.7%; Score 2637; DB 22; Length 3858;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3131 cgaattcactgcccgtctgttttacacgtctgactgggaaacccctggcgttacccaac 3190
 Db |||||||
 Qy 1187 cgaattcactgcccgtctgttttacacgtctgactgggaaacccctggcgttacccaac 1246
 Db |||||||
 Qy 3191 ttaatcccttgagcacatccccccttcgccagctggcgaatggcgctgacgagcccgca 3250
 Db |||||||
 Qy 1247 ttaatcccttgagcacatccccccttcgccagctggcgaatggcgctgacgagcccgca 1306
 Db |||||||
 Qy 3251 ccgatcccttcaccaacagttgcgcagcctgaatggcgctgacgagcccgctatt 3310
 Db |||||||
 Qy 1307 ccgatcccttcaccaacagttgcgcagcctgaatggcgctgacgagcccgctatt 1366
 Db |||||||
 Qy 3311 ttctcctaogcatctgtgcgggtatttcacaccgcataatggtgacctctcagtaacaatct 3370
 Db |||||||
 Qy 1367 ttctcctaogcatctgtgcgggtatttcacaccgcataatggtgacctctcagtaacaatct 1426
 Db |||||||

Qy 3371 gctctgatgcccgcataagtaagccagcccgacaccccgccaaacaccccgctgacgagccct 3430
 Db |||||||
 Qy 1427 gctctgatgcccgcataagtaagccagcccgacaccccgccaaacaccccgctgacgagccct 1486
 Qy 3431 gacgggctgtctgtctcccgccatccgcttacagacagcgtgtgacctctcccgagagct 3490
 Db |||||||
 Qy 1487 gacgggctgtctgtctcccgccatccgcttacagacagcgtgtgacctctcccgagagct 1546
 Qy 3491 gcatgtgtcagaggttttcacgcgtcatccacgaaacgcgcagacgaaagggccctcgtga 3550
 Db |||||||
 Qy 1547 gcatgtgtcagaggttttcacgcgtcatccacgaaacgcgcagacgaaagggccctcgtga 1606
 Qy 3551 tacgcctatttttataggttaagtgcataataatggtttcttagcgtcagtgagca 3610
 Db |||||||
 Qy 1607 tacgcctatttttataggttaagtgcataataatggtttcttagcgtcagtgagca 1666
 Qy 3611 ctcttcgggggaaatgtgcgcgaaacccctatttcttattttctaaatacaata 3670
 Db |||||||
 Qy 1667 ctcttcgggggaaatgtgcgcgaaacccctatttcttattttctaaatacaata 1726
 Qy 3671 tgcacgcgtcatgagacaataaacccctgataaaatgcttcaataatttgaataaaggaaga 3730
 Db |||||||
 Qy 1727 tgcacgcgtcatgagacaataaacccctgataaaatgcttcaataatttgaataaaggaaga 1786
 Qy 3731 gtagagatttcaaacatttcgcgtgcgcctattcccttttttgcggcattttgccttc 3790
 Db |||||||
 Qy 1787 gtagagatttcaaacatttcgcgtgcgcctattcccttttttgcggcattttgccttc 1846
 Qy 3791 ctgttttgcaccccgaaacgcgtgtgaaagttaaagatgcgtgaagatcagttgggtg 3850
 Db |||||||
 Qy 1847 ctgttttgcaccccgaaacgcgtgtgaaagttaaagatgcgtgaagatcagttgggtg 1906
 Qy 3851 cacgagtggtttacatcgactggaatctcaacagcggtaagatccttgagagttttcgc 3910
 Db |||||||
 Qy 1907 cacgagtggtttacatcgactggaatctcaacagcggtaagatccttgagagttttcgc 1966
 Qy 3911 ccgaagaacgttttcccaatgatgagcacttttaaaatctctgctatgtggcggtattat 3970
 Db |||||||
 Qy 1967 ccgaagaacgttttcccaatgatgagcacttttaaaatctctgctatgtggcggtattat 2026
 Qy 3971 cccgtattgacgcggggaacagacaaactcgcgcgcacatacacactattctcagaatgact 4030
 Db |||||||
 Qy 2027 cccgtattgacgcggggaacagacaaactcgcgcgcacatacacactattctcagaatgact 2086
 Qy 4031 tgggttgactaccagctcacagaaaagcattcttacggatggcagatgacagtaagagaat 4090
 Db |||||||
 Qy 2087 tgggttgactaccagctcacagaaaagcattctacggatggcagatgacagtaagagaat 2146
 Qy 4091 tatgcagtgctgcataaccatgataacactgcggccaacttaacttctgacaacga 4150
 Db |||||||
 Qy 2147 tatgcagtgctgcataaccatgataacactgcggccaacttaacttctgacaacga 2206
 Qy 4151 tcggagacccgaagagcgttaacccgttttttgcaacaatcggggatcatgtaactcgc 4210
 Db |||||||
 Qy 2207 tcggagacccgaagagcgttaacccgttttttgcaacaatcggggatcatgtaactcgc 2266
 Qy 4211 ttgatcgttgggaacccgagcgtgaatgaagccataaccacacgagcgtgacaccacga 4270
 Db |||||||
 Qy 2267 ttgatcgttgggaacccgagcgtgaatgaagccataaccacacgagcgtgacaccacga 2326
 Qy 4271 tgcctgtagaataaggcaaacacgttgcgaacactattaaactggggaactactactctag 4330
 Db |||||||
 Qy 2327 tgcctgtagaataaggcaaacacgttgcgaacactattaaactggggaactactactctag 2386
 Qy 4331 ctcccgccacaataataatagactggaaggagcggataaaagtgcagagaccactctgc 4390
 Db |||||||
 Qy 2387 ctcccgccacaataataatagactggaaggagcggataaaagtgcagagaccactctgc 2446
 Qy 4391 gctcggccctcccggtggtggtttattgctgataaaatctggagcccggtgagcgtgggt 4450
 Db |||||||
 Qy 2447 gctcggccctcccggtggtggtttattgctgataaaatctggagcccggtgagcgtgggt 2506
 Qy 4451 ctccgggtatcatgtcagcactgggggcagatggtaaagccctcccggtatcgtgattatct 4510

|||||
Db 2507 ctcgaggatcattcagcactcgggcccagatggtaagccctcccgatacgtagtattct 2566
QY 4511 acacgacgggagtcaggcaactatgatgaacgaataatagacagatcgcgtgagataggtg 4570
|||||
Db 2567 acacgacgggagtcaggcaactatgatgaacgaataatagacagatcgcgtgagataggtg 2626
QY 4571 cctcactgattaaagcatggtaactgtcagacccaagtttactcaataatacttttagattg 4630
Db 2627 cctcactgattaaagcatggtaactgtcagacccaagtttactcaataatacttttagattg 2686
QY 4631 atttaaaactcatttttaatttaaaagatctagtggaagatcccttttgataactca 4690
Db 2687 atttaaaactcatttttaatttaaaagatctagtggaagatcccttttgataactca 2746
QY 4691 tgaccaaatacccttaactgaggtttctccactgagcgtcagaccgccgtagaaaga 4750
Db 2747 tgaccaaatacccttaactgaggtttctccactgagcgtcagaccgccgtagaaaga 2806
QY 4751 tcaaggattctcttgagatccctttttcttgogogtaactgtcgtctgcaacaaaaa 4810
Db 2807 tcaaggattctcttgagatccctttttcttgogogtaactgtcgtctgcaacaaaaa 2866
QY 4811 aaccaccgtaccagcgtggtttgtttccggatcaagagctacaaactcttttcga 4870
Db 2867 aaccaccgtaccagcgtggtttgtttccggatcaagagctacaaactcttttcga 2926
QY 4871 agttaactggtctcagcagcagcagataccaaaactgtccctctagtgtagccgtagt 4930
Db 2927 agttaactggtctcagcagcagcagataccaaaactgtccctctagtgtagccgtagt 2986
QY 4931 tagggcaccacttcaagaactctgtagaccgccctacatacactcgtctgtaactcgt 4990
Db 2987 tagggcaccacttcaagaactctgtagaccgccctacatacactcgtctgtaactcgt 3046
QY 4991 taccagtgcctcgcagtcgcagataactcgtcttaccgggttgactcaagcgt 5050
Db 3047 taccagtgcctcgcagtcgcagataactcgtcttaccgggttgactcaagcgt 3106
QY 5051 agttaccgataaaggcgcagcgtggtgacgggggttcgtgcacacagcccagct 5110
Db 3107 agttaccgataaaggcgcagcgtggtgacgggggttcgtgcacacagcccagct 3166
QY 5111 tggagcgaacacctacacccagactgagatactacagcgtgagctatgagaaagccca 5170
Db 3167 tggagcgaacacctacacccagactgagatactacagcgtgagctatgagaaagccca 3226
QY 5171 cgtctccgaaggagaaagcgcagactatccggttaagcgcagggctcgaacaggag 5230
Db 3227 cgtctccgaaggagaaagcgcagactatccggttaagcgcagggctcgaacaggag 3286
QY 5231 agcgcagaggagcttcagggggaacgcctggtatctttatagctcgtcggtttc 5290
Db 3287 agcgcagaggagcttcagggggaacgcctggtatctttatagctcgtcggtttc 3346
QY 5291 gcaactctgaactgagcgtcagttttgtgactcgtcgaaggggcgagccctatga 5350
Db 3347 gcaactctgaactgagcgtcagttttgtgactcgtcgaaggggcgagccctatga 3406
QY 5351 aaaaacgcagcagcgccttttaccggttcttgcccttttgctgacctttgctcaca 5410
Db 3407 aaaaacgcagcagcgccttttaccggttcttgcccttttgctgacctttgctcaca 3466
QY 5411 tgtttcttctcgttattccctgattctgtgataaacctgattaccgctttgagtgag 5470
Db 3467 tgtttcttctcgttattccctgattctgtgataaacctgattaccgctttgagtgag 3526
QY 5471 ctgataacctcgcgcagcagcgaacacgcagcagcagcagtcagtgagcaggagcgg 5530
Db 3527 ctgataacctcgcgcagcagcgaacacgcagcagcagcagtcagtgagcaggagcgg 3586
QY 5531 aagagcggcccaatacgaacacccctctcccgcgcttgccgattcattaatgagct 5590
|||||

Db 3587 aagagcgcccaataacgcaaacccgcctctcccccgcgcttgccgattcaatgacgct 3646
QY 5591 ggacgacaggtttcccgactggaagcggcagtgagcgcaacgcaataatgtagtt 5650
|||||
Db 3647 ggacgacaggtttcccgactggaagcggcagtgagcgcaacgcaataatgtagtt 3706
QY 5651 agctcactcattaggcaccgccaggtttacactttatgtctccgctcgtatgtgtg 5710
Db 3707 agctcactcattaggcaccgccaggtttacactttatgtctccgctcgtatgtgtg 3766
QY 5711 gaattgtgacggataacaatttcacaggaagaaacgctatgacccatgattacgca 5767
Db 3767 gaattgtgacggataacaatttcacaggaagaaacgctatgacccatgattacgca 3823
RESULT 3
AAD09269
ID AAD09269 standard; DNA; 4960 BP.
XX
AC AAD09269;
XX
DT 12-SEP-2001 (first entry)
XX
DE pCMV-I-Cre-pA vector DNA.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW gene alteration; fusion protein; Human immunodeficiency virus;
KW HIV; pCMV-I-Cre-pA vector; ds.
XX
OS Unidentified.
XX
PN WO200149832-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-EP00060.
XX
PR 07-JAN-2000; 2000EP-0100351.
PR 10-NOV-2000; 2000EP-0124595.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Schwenk F;
XX
DR WPI; 2001-441873/47.
XX
PT Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX
PS Example; Page 62-63; 85pp; English.
XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAR peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a pCMV-I-Cre-pA vector DNA.
XX
SQ Sequence 4960 BP; 1225 A; 1213 C; 1296 G; 1226 T; 0 other;
Query Match 45.7%; Score 2637; DB 22; Length 4960;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3131 cgaattcactgcccgtcgttttacaacgctgactgggaaacccctggcgttacccaac 3190
|||||
Db 2289 cgaattcactgcccgtcgttttacaacgctgactgggaaacccctggcgttacccaac 2348
QY 3191 ttaatgccttcagcacatcccccttcccgagctggcgtaatagcgaagagcccgca 3250
|||||

|||||
Db 2349 ttaatcgcccttgagacatcccccctttcgccagctggcgtaataagcgaaagagcccgca 2408
Qy 3351 ccgtagcccttcccaacagttagcgagcctggaatggcggaatggcgctgtagcggtatt 3310
Db 2409 ccgtagcccttcccaacagttagcgagcctggaatggcggaatggcgctgtagcggtatt 2468
Qy 3311 ttctcccttaagcactctgtaggtattttcaacacgcgcatatagtgacactctcagtaacaatct 3370
Db 2469 ttctcccttaagcaactctgtaggtattttcaacacgcgcatatagtgacactctcagtaacaatct 2528
Qy 3371 gctctgtagcgcatagtagaagcagcccgacaccccgcccaaccccgctgacgcgacct 3430
Db 2529 gctctgtagcgcatagtagaagcagcccgacaccccgcccaaccccgctgacgcgacct 2588
Qy 3431 gacggggttctctcccggtacatccggtacagacagctgtgacacgctgtcccgaggct 3490
Db 2589 gacggggttctctcccggtacatccggtacagacagctgtgacacgctgtcccgaggct 2648
Qy 3491 gcatgtagcagaggttttccacggtcatccacgaaacgcgcgagacgaaagggcctctgga 3550
Db 2649 gcatgtagcagaggttttccacggtcatccacgaaacgcgcgagacgaaagggcctctgga 2708
Qy 3551 taecgctattttataggttaagtcatgataataatggtttcttagacgtcaggtggca 3610
Db 2709 taecgctattttataggttaagtcatgataataatggtttcttagacgtcaggtggca 2768
Qy 3611 cttttcggggaaatgtgcgggaaacccctatttgtttatttttcaatacatccaata 3670
Db 2769 cttttcggggaaatgtgcgggaaacccctatttgtttatttttcaatacatccaata 2828
Qy 3671 tgtagcgcgtcatagagacaataaaccttgataaaatgcttcaataatattgaaaaaggaaga 3730
Db 2829 tgtagcgcgtcatagagacaataaaccttgataaaatgcttcaataatattgaaaaaggaaga 2888
Qy 3731 gtagagattcaacatttcggtgcgccttattcccttttttgcggcattttgccttc 3790
Db 2889 gtagagattcaacatttcggtgcgccttattcccttttttgcggcattttgccttc 2948
Qy 3791 ctgttttttgcaccccgaaacgcgtgtgaaagtaaaagatgctgaagatcagttgggtg 3850
Db 2949 ctgttttttgcaccccgaaacgcgtgtgaaagtaaaagatgctgaagatcagttgggtg 3008
Qy 3851 cagcagtggtgtacatcgaactgtagtctcaacagcgtgaagatcctgagagattttcgcc 3910
Db 3009 cagcagtggtgtacatcgaactgtagtctcaacagcgtgaagatcctgagagattttcgcc 3068
Qy 3911 ccgaagaacgttttcccaatgatgagcacttttaaagttctgctatgtggcgcggtattat 3970
Db 3069 ccgaagaacgttttcccaatgatgagcacttttaaagttctgctatgtggcgcggtattat 3128
Qy 3971 cccgtattgacgcggggaagagacaactcggtcgccgcacatacactatttctcagaatgact 4030
Db 3129 cccgtattgacgcggggaagagacaactcggtcgccgcacatacactatttctcagaatgact 3188
Qy 4031 tgggttagtactaccagttcacagaaaaagcatcttaacgtaggcatgagcatgacaagaagaat 4090
Db 3189 tgggttagtactaccagttcacagaaaaagcatcttaacgtaggcatgagcatgacaagaagaat 3248
Qy 4091 tatcagtgctgcataaaccttagtgataaacactgcggccaacttactcttgacaacga 4150
Db 3249 tatcagtgctgcataaaccttagtgataaacactgcggccaacttactcttgacaacga 3308
Qy 4151 tcggaggaccgaagagtaaccgcttttttgcaacaatgggggataatgtaactcgcc 4210
Db 3309 tcggaggaccgaagagtaaccgcttttttgcaacaatgggggataatgtaactcgcc 3368
Qy 4211 ttgtagctgtgggaacccgagctgaatgaagccataccaaaaagcagcgtgacaccacga 4270
Db 3369 ttgtagctgtgggaacccgagctgaatgaagccataccaaaaagcagcgtgacaccacga 3428
Qy 4271 tgctgttagcaatggcaacaacgttgccgaacactataaactggcggaactacttactctag 4330
|||||

Db 3429 tgccgtgtagcaatggcaacaacgttgccgaacatttaactggcggaactaactactctag 3488
Qy 4331 ctcccggaacaataatagactggatggaggcggaataaagttagcaggaacactcttcgc 4390
Db 3489 ctcccggaacaataatagactggatggaggcggaataaagttagcaggaacactcttcgc 3548
Qy 4391 gctcggcccttccggctggtggtttattgctgataaaatctggagccggtgagcgtgggt 4450
Db 3549 gctcggcccttccggctggtggtttattgctgataaaatctggagccggtgagcgtgggt 3608
Qy 4451 ctccgggtatcatctgcagcactggggccagatggtaaacctcccgatcgtagttatct 4510
Db 3609 ctccgggtatcatctgcagcactggggccagatggtaaacctcccgatcgtagttatct 3668
Qy 4511 acacgacgggagtcaggcaactatggatgaacgaataagacagatcgtcagataggtg 4570
Db 3669 acacgacgggagtcaggcaactatggatgaacgaataagacagatcgtcagataggtg 3728
Qy 4571 cctcactgattaaagcattggtaactgtcagaccgaagttaactcatatatacttttagattg 4630
Db 3729 cctcactgattaaagcattggtaactgtcagaccgaagttaactcatatatacttttagattg 3788
Qy 4631 atttaaaacttcaatttttaatttaaaagatctaggtgaagatccttttgataatctca 4690
Db 3789 atttaaaacttcaatttttaatttaaaagatctaggtgaagatccttttgataatctca 3848
Qy 4691 tgacaaaatacccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaaga 4750
Db 3849 tgacaaaatacccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaaga 3908
Qy 4751 tcaaaaggatcttttgagatcctttttctgcgcgttaactgtcgtgctgctgaacaaaaa 4810
Db 3909 tcaaaaggatcttttgagatcctttttctgcgcgttaactgtcgtgctgctgaacaaaaa 3968
Qy 4811 aaccacccgtaccagcgtggtttggttgcgggatacagagctacacactcttttccga 4870
Db 3969 aaccacccgtaccagcgtggtttggttgcgggatacagagctacacactcttttccga 4028
Qy 4871 aggttaactggtctcagcagagcgcagatacccaataactgctctcttagtagcgtgagt 4930
Db 4029 aggttaactggtctcagcagagcgcagatacccaataactgctctcttagtagcgtgagt 4088
Qy 4931 taggcacacacttcaagaactctgtagcacgcgtcacaatacctcgtctgtctgtaactctgt 4990
Db 4089 taggcacacacttcaagaactctgtagcacgcgtcacaatacctcgtctgtctgtaactctgt 4148
Qy 4991 taccagtggtcgtgcgcagtgccgataagtgctcttaccgggttgagactcaagacgat 5050
Db 4149 taccagtggtcgtgcgcagtgccgataagtgctcttaccgggttgagactcaagacgat 4208
Qy 5051 agttaccgggataaaggcgcagcgtggtggtgaaacggggggttcgtgcacacagccagct 5110
Db 4209 agttaccgggataaaggcgcagcgtggtggtgaaacggggggttcgtgcacacagccagct 4268
Qy 5111 tggagcgaacgacactacacccgaactgagataccttacagcgtgagctatgaaaaagcgca 5170
Db 4269 tggagcgaacgacactacacccgaactgagataccttacagcgtgagctatgaaaaagcgca 4328
Qy 5171 cgcttcccgaaaggagaaaggcgcaggtatccggttaagcgcaggggtcggaacagagag 5230
Db 4329 cgcttcccgaaaggagaaaggcgcaggtatccggttaagcgcaggggtcggaacagagag 4388
Qy 5231 agcgacgagggagcttccagggggaacgctctgtatctttatagtcctcgtcggttttc 5290
Db 4389 agcgacgagggagcttccagggggaacgctctgtatctttatagtcctcgtcggttttc 4448
Qy 5291 gccacctgactgagcgtcgatttttgatgctcgtcagggggggcggaacccatagga 5350
Db 4449 gccacctgactgagcgtcgatttttgatgctcgtcagggggggcggaacccatagga 4508
Qy 5351 aaaacgcagcaacgcgaccttttaccggttctcgtccttttgcgcttttgcctcaca 5410
Db 4509 aaaacgcagcaacgcgaccttttaccggttctcgtccttttgcgcttttgcctcaca 4568


```
Db 5621 tatgcagtgctgccataaaccaatgagtgataaacactgctgcggccaacttaacttacttcttgacaacga 5680
Qy 4151 tcgagagaccgaagagagtaaccgcttttttgcacaacatggtgggagatcatgttaactgcc 4210
Db 5681 tcgagagaccgaagagagtaaccgcttttttgcacaacatggtgggagatcatgttaactgcc 5740
Qy 4211 ttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacacga 4270
Db 5741 ttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacacga 5800
Qy 4271 tgctctgtagcaatggcaacaacgttgcgaacactatttaacttgcgaacactaacttaactctag 4330
Db 5801 tgctctgtagcaatggcaacaacgttgcgaacactatttaacttgcgaacactaacttaactctag 5860
Qy 4331 ctccccgcaacaattaatagactgga tggaggoggataaaagtgtgcaggaaccaactttcgc 4390
Db 5861 ctccccgcaacaattaatagactgga tggaggoggataaaagtgtgcaggaaccaactttcgc 5920
Qy 4391 gctcggcccttccgctgctggtttattctgtgataaatctggagccggtgagcgtgggt 4450
Db 5921 gctcggcccttccgctgctggtttattctgtgataaatctggagccggtgagcgtgggt 5980
Qy 4451 ctgcggttatcatgcaacactggggccagatggttaagccctcccgatcgtagtattct 4510
Db 5981 ctgcggttatcatgcaacactggggccagatggttaagccctcccgatcgtagtattct 6040
Qy 4511 acagacggggagtcaggaactatggtatgaacgaaa tagacagatcgtgagatagtg 4570
Db 6041 acagacggggagtcaggaactatggtatgaacgaaa tagacagatcgtgagatagtg 6100
Qy 4571 cctcaactgattaaacattgtaactgtcagaccaagtttaactcatataacttttagattg 4630
Db 6101 cctcaactgattaaacattgtaactgtcagaccaagtttaactcatataacttttagattg 6160
Qy 4631 atttaaaacttcaattttaaattaaaagatcctagtgaaagatcccttttggataatctca 4690
Db 6161 atttaaaacttcaattttaaattaaaagatcctagtgaaagatcccttttggataatctca 6220
Qy 4691 tgacaaatacccttaacgtgagtttcttcactgagcgtcagaccccgtagaaaaa 4750
Db 6221 tgacaaatacccttaacgtgagtttcttcactgagcgtcagaccccgtagaaaaa 6280
Qy 4751 tcaaaagatctcttgagatcctttttctgcgtaactgtcgtctgcaacaaaaa 4810
Db 6281 tcaaaagatctcttgagatcctttttctgcgtaactgtcgtctgcaacaaaaa 6340
Qy 4811 aaccacgctaccagcgtgtgtttgttgcgggatacaagagctaccaactcttttccga 4870
Db 6341 aaccacgctaccagcgtgtgtttgttgcgggatacaagagctaccaactcttttccga 6400
Qy 4871 aggttaactggttcagcagagcgagataccaaaatactgtccttctagtgtagccgtagt 4930
Db 6401 aggttaactggttcagcagagcgagataccaaaatactgtccttctagtgtagccgtagt 6460
Qy 4931 tagccaccacttcaagaaactctgtagcacgcgcctacataactcgtcgtctgtaactcgt 4990
Db 6461 tagccaccacttcaagaaactctgtagcacgcgcctacataactcgtcgtctgtaactcgt 6520
Qy 4991 taccagtggtcgtgccagtgccgataagtcgtgtcttaaccgggttggaactcaagaacgat 5050
Db 6521 taccagtggtcgtgccagtgccgataagtcgtgtcttaaccgggttggaactcaagaacgat 6580
Qy 5051 agttaccogga taagggcagcggctcggtgtaacggggggttcgtgcaacagccagct 5110
Db 6581 agttaccogga taagggcagcggctcggtgtaacggggggttcgtgcaacagccagct 6640
Qy 5111 tggagcgaacacctaaccgaactgaatacctatacagcgtgagctatgaaagcgcca 5170
Db 6641 tggagcgaacacctaaccgaactgaatacctatacagcgtgagctatgaaagcgcca 6700
Qy 5171 cgcctcccgaaagggagaaagcgacaggtatccggttaagcgaggtcgtggaacagag 5230
Db 6701 cgcctcccgaaagggagaaagcgacaggtatccggttaagcgaggtcgtggaacagag 6760
```

```
Qy 5231 agcgacagaggagagcttccaggggaaacgcctgggtatatttattatgtctcgtcggttcc 5290
Db 6761 agcgacagaggagagcttccaggggaaacgcctgggtatatttattatgtctcgtcggttcc 6820
Qy 5291 gccacctctgactgagcgtcgatttttgcagtcgtcagggggggagacccatgga 5350
Db 6821 gccacctctgactgagcgtcgatttttgcagtcgtcagggggggagacccatgga 6880
Qy 5351 aaaaacgcagcaacgcggcctttttacagttcctcggcttttgcggccttttgcctaca 5410
Db 6881 aaaaacgcagcaacgcggcctttttacagttcctcggcttttgcggccttttgcctaca 6940
Qy 5411 tgttttctcgtcgttatccccctgattctgtgtgataacogtattaccgcctttgagtgag 5470
Db 6941 tgttttctcgtcgttatccccctgattctgtgtgataacogtattaccgcctttgagtgag 7000
Qy 5471 ctgataccgcctcgcgcagcgcgaacgcgacgcgagcgcagcgcagtcagtcgagcgaagcgg 5530
Db 7001 ctgataccgcctcgcgcagcgcgaacgcgacgcgagcgcgagtcagtcgagcgaagcgg 7060
Qy 5531 aagagcgcccaatacagcaaacgcctctcccgcgcttggccgattcatttaatgcagct 5590
Db 7061 aagagcgcccaatacagcaaacgcctctcccgcgcttggccgattcatttaatgcagct 7120
Qy 5591 ggcacgacaggtttcccgcactggaagcgggcagtcgagcgaacgcaatataatgtgatt 5650
Db 7121 ggcacgacaggtttcccgcactggaagcgggcagtcgagcgaacgcaatataatgtgatt 7180
Qy 5651 agctcaactcattagggcacccacaggtttacactttatgcttccggtcgtatgtgtgtg 5710
Db 7181 agctcaactcattagggcacccacaggtttacactttatgcttccggtcgtatgtgtgtg 7240
Qy 5711 gaattgtgagcggatacaaatcttcacaggaacagcgttatgacctatgacctacgcga 5767
Db 7241 gaattgtgagcggatacaaatcttcacaggaacagcgttatgacctatgacctacgcga 7297
```

RESULT 5
AAD04946/C

ID AAD04946 standard; DNA; 8062 BP.

XX AAD04946;

XX 17-JUL-2001 (first entry)

XX Plasmid pRK76 used to test Cre recombinase mediated inversion.

XX DE Gene trapping construct; conditional mutation; unidirectional inversion;
XX KW recombinase recognition sequence; RRS; disruption cassette;
XX KW selection cassette; transgenic organism; pRK76; Cre recombinase; ds.

XX OS Chimeric - Adenovirus.

XX OS Chimeric - ECMV virus.

XX OS Chimeric - Unidentified.

XX PN WO200129208-A1.

XX XX 26-APR-2001.

XX PF 16-OCT-2000; 200WO-EP10162.

XX PR 16-OCT-1999; 99EP-0120592.

XX PR 27-OCT-1999; 99US-0162016.

XX XX (ARTE-) ARTEMIS PHARM GMBH.

XX PA (FRAN-) FRANKEN BIOTECHNOLOGIE AG.

XX PI Kuehn R, Von Melchner H, Altschmid J;

XX DR WPT; 2001-308486/32.

XX PT New gene trapping construct capable of causing conditional mutations in

PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX Example 3; Page 70-73; 78pp; English.

XX The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is pRK76 vector, which is used to test Cre
CC recombinase mediated inversion, which is related to the invention.
XX
SQ Sequence 8062 BP; 1960 A; 2036 C; 2121 G; 1945 T; 0 other;

Query Match: 45.7%; Score 2637; DB 22; Length 8062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggcgtcgcttttacaacgctgactggaacacccctgcgtttaccacaac 3190
DB 8045 CGAATTCATGGCGCGTGGTTCACACGCTGCTGACTGGGAAACCCCTGGCGTTACCCCAAC 7986
QY 3191 ttaatcgcccttcgacacatccccccttcgcagctggcgtaagcgaagagcccgca 3250
DB 7985 TTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAGAGGCCGCA 7926
QY 3251 cgaatcgcccttcgaacagcttcgcagcctgaatggcgaatggcgccctgagcgtatt 3310
DB 7925 CGATCGCCCTTCCCAACAGTTGCGCAGCGCTGAATGGCGAATGGCGCCTGATCGCGTATT 7866
QY 3311 ttctcctacgcactgtcggttatctcacaccgcataagtgacacttcagtaaatct 3370
DB 7865 TTCCTCTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCT 7806
QY 3371 gctctgatccgcatagttaagccagcccgacaccccgcaacacccgctgacgcgccct 3430
DB 7805 GCTCTGATGCCGATAGTTAAGCCAGCCCGCACACCCGCCCAACACCCGCTGACGCCCT 7746
QY 3431 gacggcttgctgtcccggaatccgcttacagacaagctgtgacgcttcocggagct 3490
DB 7745 GACGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCT 7686
QY 3491 gcatgtgcagaggtttccacgctatccacgaacacgcgcgacgaagaaaggccctgta 3550
DB 7685 GCATGTGCAGAGGTTTCCACGCTATACCGAACCGCGAGACGAAGGGCCTCGTA 7626
QY 3551 tacgcctattttataggttaattgcatgataataatgggtttcttagacgtcaggtggca 3610
DB 7625 TAGCGCTATTTTATAGGTTAATGTCATGATAATAATGTTTCTTAGACGTCAGGTGGCA 7566
QY 3611 cttttcggggaatgtgcgcgaacccctatttgtttatttttcttaataacattcaata 3670
DB 7565 CTTTTCGGGGAATGTGCGCGAACCCTTATTTGTTTATTTTCTAAATACATTCAATA 7506
QY 3671 tctatccgctcatgagacaataaccctgataaatgctcaataattcaaaaagaa 3730
DB 7505 TGTATCCGCTCATGAGACATAACCTGATAATGCTTCAATAATATTGAAAAGGAAGA 7446
QY 3731 gtagagattcaaacatttcgctgcgccttatcccttttttggcgcattttgccttc 3790
DB 7445 GTATGAGTATCAACATTTCCGCTGCGCCTTATTCCTTTTTTTGCGGCATTTTGCCTTC 7386

QY 3791 ctgttttctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3850
DB 7385 CTGTTTTTCTCACCAGAAACGCTGGTGAAGATAAAGATGCTGAAGATCAGTTGGGTG 7326
QY 3851 cacgagtggtttacatcgaactggatctcaacagcggtgaagatccctgagagtttcgccc 3910
DB 7325 CACGAGTGGGTTACATCGAACTGATCTCAACAGCGGTAGATCCTTGAGAGTTTCGCC 7266
QY 3911 ccgaagaacggttttccaatgatgagcaccttttaaaagttctgctatgtggcggtattat 3970
DB 7265 CCGAAGAACGCTTTTCCAATGATGAGCACCTTTAAAGTTCTGCTATGTGGCGCGTATTAT 7206
QY 3971 ccgattatgacgcgggcaagagaacactcggtcgcgacataactattctcagaatgact 4030
DB 7205 CCGGTATTGACGCGGCAAGACAACTCGGTGCGCGCATACACTATTCTCAGAATGACT 7146
QY 4031 tggttgactactcaccagtcacagaaaagcatcttcaggtatgcatgacagtaagaagt 4090
DB 7145 TGGTTGAGTACTCACCAGTCCAGAAAAGCATCTTACGGATGCGATGACAGTAAGAGAT 7086
QY 4091 tatcgagtgctgcataaacatcagtgatgataaacactcgggccaaacttactctcgacaacga 4150
DB 7085 TATGCAGTGTGTCATTAACCATGAGTGATAACACTGCGGCCAACTTACTTCTCACAACGA 7026
QY 4151 tcggagacggagagactaacgcgttttttgcacaaatcgggggacatgtaactcgcc 4210
DB 7025 TCGGAGACCGGAAGGAGCTAAACGCTTTTTCACAACTGGGGGATCATGTAACTCGCC 6966
QY 4211 ttgatcgttgggaacggagctgaatgaagccatacaaacacgacgagcgtgacacacga 4270
DB 6965 TTGATCGTTGGGAACCGGAGCTGAATGAGCCATACCAAAACGACGAGCGTGACACCACGA 6906
QY 4271 tgctctagcaatggcaacaacgcttcgcaaaactattaaactggcgaactacttactctag 4330
DB 6905 TGCTCTAGCAATGGCAACAACCTTGCGCAACTATTAACTGGCGAACTACTTACTCTAG 6846
QY 4331 ctcccggaacaacatataagactggagcggaataaagtgtcaggacacactctcgc 4390
DB 6845 CTTCGCGCAACAATTAATAGACTGGAGCGCGATAAAGTTGCAGGACCACTTCTGC 6786
QY 4391 gctcgcccttcgctgctggtttatctgataaactggagcggcgagcgtgggt 4450
DB 6785 GCTCGGCCCTTCGCGCTGCTGCTGTTTATGCTGATTAATCTGGAGCGGTGACGCGGGT 6726
QY 4451 ctgcggtatcatctgcagcactggggccagatggtgaagccctccctctatcgtatgtatct 4510
DB 6725 CTGCGGTATCATTTGACGACTGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6666
QY 4511 acagcggggagtcaggcaactatggatgaacgaataagacagatcgctgagatagggtg 4570
DB 6665 ACACGCGGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTG 6606
QY 4571 cctcactgattgaagcattgttaactgtcagacaagtttactcatatatactttagattg 4630
DB 6605 CTTCACTGATTAGCAATGGTAACCTGTACAGCCCAAGTTTACTCATATATATCTTTAGATTG 6546
QY 4631 atttaaaacttcaatttttaatttaaagatctagtggaagatccctttttgataactca 4690
DB 6545 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGTGGAAGATCCCTTTTGTGATACTCA 6486
QY 4691 tgacaaaaatcccttaacgtgagctttcgttcacactgagcgtcagaccccgtagaagaaga 4750
DB 6485 TGACCAAAAATCCCTTAACGTGAGTTTTTTCGTTCCACTGAGCGTCAGACCCCGTGAAGAAGA 6426
QY 4751 tcaagatctcttgagatcccttttctcgcgtaactctgctgctctgcaacaaaaa 4810
DB 6425 TCAAGAGATCTTCTTGAGATCCCTTTTTTCTGCGCGTAACTCTGCTCTGTGCAACAAAAA 6366
QY 4811 aaccacccgtaccagcggtgtgtttgttcgggatacaagagctaccacactcttttccga 4870
DB 6365 AACCAACCGCTACCAAGCGGTGTTGTTGGCGGATCAAGAGCTACCAACTCTTTTCCGA 6306
QY 4871 aggtaactggcttcagcagagcgacagatacccaaaactgctcctctctagtagtagccgtagt 4930


```
Db 6305 AGGTAACCTGGCTTCAGCAGACGGCAGATACCAATACTGTCTCTTAGTAGCCGTAGT 6246
Qy 4931 taggcaccactcaagaactctgtagcacgcgcctacatacctcgtcgtctaaacctgt 4990
Db 6245 TAGGCCACCCTCAAGAACCTCTAGCACCAGCCTACATACCTCGCTCTGCTAATCTGT 6186
Qy 4991 taccagtggtcgtcgaagtgcgcataaagtcgtcttaccggttggaactcaagaagat 5050
Db 6185 TACCAGTGGCTGCTGCCAGTGGCAGTAAGTCTGTCTTACCGGGTTGACATCAAGACGAT 6126
Qy 5051 agttaccgataaagcgcagcgttcgagctaaacgggggttcctgcacacagcccaact 5110
Db 6125 AGTTACCGGATAAGCGCAGCGCTCGGCTGAACGGGGGGTTCGTGCACACAGCCACGT 6066
Qy 5111 tggagcgaacacacacacacacacacacacacacacacacacacacacacacacacac 5170
Db 6065 TGGAGCAACACACCTACACCACTAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 6006
Qy 5171 cgttcccgaaaggagaagcgcagaggtatccggttaagcgcagggcagggcgaacaggag 5230
Db 6005 CGCTTCCGGAAGGAGAAAGCGCAGAGGTATCCGGTAAGCGCGCAGGTCGGAACAGGAG 5946
Qy 5231 agcgcagagggagcttcacaggggaaacgcctgttatcttatagtcctcgtcggttcc 5290
Db 5945 AGCGCAGGAGGAGCTTCCACGGGGAAACGCGCTGGTATCTTTTATAGTCTCTCGGGTTTC 5886
Qy 5291 gccacctctgactgagcgtcgatttttctgtagctcgtcgtaagggggcggagcctatgga 5350
Db 5885 GCCACCTCTGACTTGAGCGTCAATTTTGTGATGCTGTCACAGGGGGGAGCCTATGGA 5826
Qy 5351 aaacgcagcaacgcgcgttttttaaggttcctcgttcctcgttcctcgttcctcgttc 5410
Db 5825 AAAACGCCAGCAACCGCGCCCTTTTACGGTTCTCCTGGCCCTTTTGTGGCCCTTTTGTCTACA 5766
Qy 5411 tgttcttctcgttataccctcgtattctgttgataaacctattaccgcctttgagtgag 5470
Db 5765 TGTCTTTCTCTGCGTTATCCCTGATCTGTGGATAACCGTATTAACCGCCTTGAAGTGAG 5706
Qy 5471 ctgatacgcctcgcgcagcgaacgcgcagcgcagcgcagcgcagcgcagcgcagcgcg 5530
Db 5705 CTGATACCGCTTCGGCGCAGCGCAACAGCAGCGCAGCGCAGCTCAGTGACGAGGAAGCGG 5646
Qy 5531 aagagcgcacaaacgcgcctctcccgcgcttcctccgcgcgttgcgattcattaatcagct 5590
Db 5645 AAGAGCGCCCAATAGCAAAACCGCTCTCCCGCGCGCTTGGCCGATTATTATGACAGCT 5586
Qy 5591 ggcacgacaggtttcccgactggaagcggcagtgagcgaacgaattaatgtgagtt 5650
Db 5585 GGCACGACAGCTTCCCGACTTGAAGAGCGGCGCAGTGACGCAACGCAATTAATGTGAGTT 5526
Qy 5651 agctcaactcattaggacccacagcctttacactttatgcttcgggctcgtatgttggtg 5710
Db 5525 AGCTCACATTAATAGCAGCCCGAGGCTTTACACTTTATGCTTCGGCTCGTATGTTGTGTG 5466
Qy 5711 gaattgtgagcggataaactttcacacagaaacagcgtatgacatgattacgcca 5767
Db 5465 GAATGTGACGGATACAAATTTACACAGGAACAGCTATGACCATGATTACGCCA 5409
```

```
RESULT 6
ID AAD04945/c
XX AAD04945 standard; DNA; 8153 BP.
AC AAD04945;
XX
DT 17-JUL-2001 (first entry)
XX
DE Plasmid prk74 used to test Cre recombinase mediated inversion.
XX
KW Gene trapping construct; conditional mutation; unidirectional inversion;
KW recombinase recognition sequence; RRS; disruption cassette;
KW selection cassette; transgenic organism; prk74; Cre recombinase; ds.
```

```
XX Chimeric - Adenovirus.
OS Chimeric - Unidentified.
OS Chimeric - ECMV virus.
PN WO200129208-A1.
XX
PD 26-APR-2001.
XX
PF 16-OCT-2000; 2000WO-EP10162.
XX
PR 16-OCT-1999; 99EP-0120592.
PR 27-OCT-1999; 99US-0162016.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
PA (FRAN-) FRANKEN BIOTECHNOLOGIE AG.
PI Kuehn R, Von Melchener H, Altschmied J;
XX
DR WPI; 2001-308486/32.
XX
PT New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped
XX
FS Example 3; Page 68-70; 78pp; English.
XX
CC The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is prk74 vector, which is used to test Cre
CC recombinase mediated inversion, which is related to the invention.
XX
SQ Sequence 8153 BP; 1976 A; 2133 C; 2067 G; 1977 T; 0 other;
```

Query Match 45.7%; Score 2637; DB 22; Length 8153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 3131 cgaattcactggccgtcgttttacaaacgtcgtgactgggaaacccctggcgttacccaac 3190
Db 8136 CGAATTACTGGCCGCTGTTTACAAACGTCGTGACTGGGAAACCTGGCGTTACCCAAC 8077
Qy 3191 ttaatcgctcgcagcacatcccccttcgccagctggcgtaataagcgaagagccgcga 3250
Db 8076 TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAATAGCGAAGGCCGCA 8017
Qy 3251 ccgatcgcccttcccaacagttgcgcagcgtgaatggcgaatggcgccctgacggtatt 3310
Db 8016 CCGATCGCCCTTCCCAACAGTTGGCAGCCTGAATGGCGAATGGCGCCTGATGCCGTATT 7957
Qy 3311 ttctccttacgcatctgtgcggtatttcacaccgcgcatatgggtgcactctcagtaacaat 3370
Db 7956 TTCTCTTACGCATCTGTGGGTATTTCACACCGCATATGGTGACTCTCAGTACAATCT 7897
Qy 3371 gctctgtgcgcgcatagttaagccagccccgcaccccccaacccccgcgtgcgcgcct 3430
Db 7896 GCTCTGTATGCGCATAGTTAAGCCAGCCGACACCCGCAACACCCGCTGACGGCCCT 7837
Qy 3431 gacgggctgtcgtcccgccatccgttacagacagcgttgaccatcctcgggagct 3490
```


Qy 5651 agctcaactcattagaccacccagcgtttacactttatgtcttcctcggcgtatgtgtgtg 5710
|||||
Db 5616 AGCTCACTCATAGGACACCCAGGCTTTACACTTTATGCTTCGGGCTCGTATGTTGTG 5557
|||||
Qy 5711 gaattgtgagcgataaacaatttcacacagaaacagctatgacatattacgcc 5767
|||||
Db 5556 GAATTGTGAGCGGATACAAATTTACACAGGAACAGCTATGACCATGATTACGCCA 5500
|||||

RESULT 7

AAZ20086
ID AAZ20086 standard; DNA; 8854 BP.
XX
AC AAZ20086;
XX
DT 05-JAN-2000 (first entry)
XX
DE DNA encoding chimeric cryIB insecticide HyFLIB.
XX
KW Insecticide; HyFLIB; crystal protein; delta-endotoxin; toxin;
KW cryIB; cryIA(b); maize; transgenic plant; European corn borer;
KW Ostrinia nubilalis; entomocide; crop protection; biological control;
KW pCIB5520; ss.
XX
OS Bacillus thuringiensis.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 11..3694
FT CDS /*tag= a
FT /*product= "HyFLIB"
XX

PN W09950293-A1.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-EP02175.

XX 01-APR-1998; 98US-0053549.

XX (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX Desai NM;

XX WPI; 1999-601323/51.

DR P-PSDB; AAY31990.

XX Chimeric insecticidal protein comprising cryIB and cryIA(b) portions

PS Claim 3; Page 34-42; 85pp; English.

XX This is the nucleotide sequence of a synthetic gene that codes for
CC a chimeric insecticidal protein, termed hyFLIB (see AAY31990), of 1227
CC amino acids composed of a cryIB core N-terminal toxin portion of
CC 845 amino acids and a cryIA(b) C-terminal protoxin portion of 382
CC amino acids. The synthetic gene is based on the cryIB and cryIA(b)
CC genes of *Bacillus thuringiensis*, with codon usage altered to
CC increase expression in plants, particularly maize. The synthetic
CC gene is contained within plasmid pCIB5520 (NRRL B-21572) under
CC control of the PEPC promoter. When the chimeric insecticidal
CC protein gene is expressed in transgenic maize from both PEPC and
CC pith promoters, insecticidal activity is observed against European
CC corn borer (*Ostrinia nubilalis*). Recombinant microbial strains
CC transformed with the hyFLIB gene can be used in endomocidal
CC formulations for the biological control of Lepidopteran pests.

XX Sequence 8854 BP; 2202 A; 2580 C; 2189 G; 1883 T; 0 other;

Query Match

Beat Local Similarity 45.7%; Score 2637; DB 20; Length 8854;

Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3131 cgaattcaactgcccgtcgcttttacaacgtctgactgggaacacctggcgttacccaac 3190
|||||
Db 3892 cgaattcaactgcccgtcgcttttacaacgtctgactgggaacacctggcgttacccaac 3951
|||||
Qy 3191 ttaatcgcttgcagcacatccccctttccgccagctggcgtaataagcgaagagcccgca 3250
|||||
Db 3952 ttaatcgcttgcagcacatccccctttccgccagctggcgtaataagcgaagagcccgca 4011
|||||
Qy 3251 ccgatacgccctcccaacagttgcgacgctgaatggcgaaatggcgctgacggtatt 3310
|||||
Db 4012 ccgatacgccctcccaacagttgcgacgctgaatggcgaaatggcgctgacggtatt 4071
|||||
Qy 3311 ttctcttacgcatctgtcggttatttcacacgcgcatatggtgcactctcagtaaatct 3370
|||||
Db 4072 ttctcttacgcatctgtcggttatttcacacgcgcatatggtgcactctcagtaaatct 4131
|||||
Qy 3371 gctctgatgcgcatagtttaagccagccgcacacccccaaccccgctgacgcccct 3430
|||||
Db 4132 gctctgatgcgcatagtttaagccagccgcacacccccaaccccgctgacgcccct 4191
|||||
Qy 3431 gacgggctgtctgtcccgccatccgcttacagacaagctgtgacctctccggagct 3490
|||||
Db 4192 gacgggctgtctgtcccgccatccgcttacagacaagctgtgacctctccggagct 4251
|||||
Qy 3491 gcatgtctagaggttttcacgcgtcatcccgaaacgcgcagacgaagggcctctgta 3550
|||||
Db 4252 gcatgtctagaggttttcacgcgtcatcccgaaacgcgcagacgaagggcctctgta 4311
|||||
Qy 3551 taacgctattttataggttaagtcatgataataatagtttcttagacgtcagtgcca 3610
|||||
Db 4312 taacgctattttataggttaagtcatgataataatagtttcttagacgtcagtgcca 4371
|||||
Qy 3611 ctttcggggaaatgtgcgcgaaacccctatttcttttctaaatcacaatcaca 3670
|||||
Db 4372 ctttcggggaaatgtgcgcgaaacccctatttcttttctaaatcacaatcaca 4431
|||||
Qy 3671 tgtatccgctcatgagacaataaaccttgataaatgcttcaataattgaaaaaagaaga 3730
|||||
Db 4432 tgtatccgctcatgagacaataaaccttgataaatgcttcaataattgaaaaaagaaga 4491
|||||
Qy 3731 gtatgagtattcaacatttcgctgcgccttattcccttttttgcgcatcttgccttc 3790
|||||
Db 4492 gtatgagtattcaacatttcgctgcgccttattcccttttttgcgcatcttgccttc 4551
|||||
Qy 3791 ctgttttgcacccagaaacgcgtgtgaaagttaaagatgctgaagatcagttgggtg 3850
|||||
Db 4552 ctgttttgcacccagaaacgcgtgtgaaagttaaagatgctgaagatcagttgggtg 4611
|||||
Qy 3851 cagcagtggtttacatcgaaactggatctcaacagcggtaagatccttgagagtttcgcc 3910
|||||
Db 4612 cagcagtggtttacatcgaaactggatctcaacagcggtaagatccttgagagtttcgcc 4671
|||||
Qy 3911 ccgaagaacggttttccaatgatgagcaacttttaaaagtctgctatgtggcggtattat 3970
|||||
Db 4672 ccgaagaacggttttccaatgatgagcaacttttaaaagtctgctatgtggcggtattat 4731
|||||
Qy 3971 cccgtattgacgcggcgcaagagcaactcgtgcgcgcatacacatattctcagatgact 4030
|||||
Db 4732 cccgtattgacgcggcgcaagagcaactcgtgcgcgcatacacatattctcagatgact 4791
|||||
Qy 4031 tggttgagtactcaccagtcacagaaaaagcatcttacggatggcgatgacagtaagaagt 4090
|||||
Db 4792 tggttgagtactcaccagtcacagaaaaagcatcttacggatggcgatgacagtaagaagt 4851
|||||
Qy 4091 tatgcagtgctgcataaccatgataacacactgcggccaacttacttctgacaacga 4150
|||||
Db 4852 tatgcagtgctgcataaccatgataacacactgcggccaacttacttctgacaacga 4911
|||||
Qy 4151 tcggaggaccgaagagcgttaaccgcttttttgcacaacatggggggtcatgttaactcgcc 4210
|||||
Db 4912 tcggaggaccgaagagcgttaaccgcttttttgcacaacatggggggtcatgttaactcgcc 4971
|||||

```
QY 4211 ttgatcgttgggaacggagctgaatgaagccatacaacacgacgagcgtgacacacga 4270
|||||
Db 4972 ttgatcgttgggaacggagctgaatgaagccatacaacacgacgagcgtgacacacga 5031
|||||
QY 4271 tgctgtagcaatggcaacacgttcgcaaacattataactgcgcgaactactactctag 4330
|||||
Db 5032 tgctgtagcaatggcaacacgttcgcaaacattataactgcgcgaactactactctag 5091
|||||
QY 4331 cttcccgcccaacaattaatagactggtgagggcgataaaattgacagaccactctcgc 4390
|||||
Db 5092 cttcccgcccaacaattaatagactggtgagggcgataaaattgacagaccactctcgc 5151
|||||
QY 4391 gctcgcccttcggctggtggtttatctgctgataaaatcgagccggtgaagcgtgggt 4450
|||||
Db 5152 gctcgcccttcggctggtggtttatctgctgataaaatcgagccggtgaagcgtgggt 5211
|||||
QY 4451 ctgcggttatctgcagcaactggcccgatggtgaagccctcccgatatcgtagttatct 4510
|||||
Db 5212 ctgcggttatctgcagcaactggcccgatggtgaagccctcccgatatcgtagttatct 5271
|||||
QY 4511 acacgacggggagtcaggcaactatggtgaacgaaatagacagatcgctgagataggtg 4570
|||||
Db 5272 acacgacggggagtcaggcaactatggtgaacgaaatagacagatcgctgagataggtg 5331
|||||
QY 4571 cctcaactgattaagcattggtaactgcagaccgaatttactcaataatactttagattg 4630
|||||
Db 5332 cctcaactgattaagcattggtaactgcagaccgaatttactcaataatactttagattg 5391
|||||
QY 4631 atttaaaactctattttaaattaaagagatctaggtgaagatcctttttgataactcga 4690
|||||
Db 5392 atttaaaactctattttaaattaaagagatctaggtgaagatcctttttgataactcga 5451
|||||
QY 4691 tgacaaaaatcccttaactgagtttcttccactgagcgtcagaccctcgtagaaaaa 4750
|||||
Db 5452 tgacaaaaatcccttaactgagtttcttccactgagcgtcagaccctcgtagaaaaa 5511
|||||
QY 4751 tcaaaagatctcttgagatccctttttctgcgcgtaatactctgcttgcaaaaacaaa 4810
|||||
Db 5512 tcaaaagatctcttgagatccctttttctgcgcgtaatactctgcttgcaaaaacaaa 5571
|||||
QY 4811 aaccacgctaccagcgttggtttgttcccggaatacaagatcaacaaactctttccga 4870
|||||
Db 5572 aaccacgctaccagcgttggtttgttcccggaatacaagatcaacaaactctttccga 5631
|||||
QY 4871 agtgaactggcttcagcagcagcagataccaaaatactgctcttctagtgtagccgtagt 4930
|||||
Db 5632 agtgaactggcttcagcagcagcagataccaaaatactgctcttctagtgtagccgtagt 5691
|||||
QY 4931 taggccacacttcaagaactctgtagaccgcctacatacctcgtctctgctaactcgt 4990
|||||
Db 5692 taggccacacttcaagaactctgtagaccgcctacatacctcgtctctgctaactcgt 5751
|||||
QY 4991 taccagtggctgctgcagtgccgataagtctgtcttaccgggttggactcaagacgat 5050
|||||
Db 5752 taccagtggctgctgcagtgccgataagtctgtcttaccgggttggactcaagacgat 5811
|||||
QY 5051 agttaccgdataaaggcagcagcgtgggtgaacgggggttctgtgcacacagccagct 5110
|||||
Db 5812 agttaccgdataaaggcagcagcgtgggtgaacgggggttctgtgcacacagccagct 5871
|||||
QY 5111 tggagcgaacgacctacacgaaactagatacctacagcgtgagctatagaaaaagccga 5170
|||||
Db 5872 tggagcgaacgacctacacgaaactagatacctacagcgtgagctatagaaaaagccga 5931
|||||
QY 5171 cgcctcccgaaaggagaaaggcgcagagatccggtgaagcgcagcagcgtcggaacagag 5230
|||||
Db 5932 cgcctcccgaaaggagaaaggcgcagagatccggtgaagcgcagcagcgtcggaacagag 5991
|||||
QY 5231 agcgcacgagggagcttccagggggaaacgcctggttatctttatagctcctgctgggttc 5290
|||||
Db 5992 agcgcacgagggagcttccagggggaaacgcctggttatctttatagctcctgctgggttc 6051
|||||
QY 5291 gccacctcgaactgagcgtcgatctttttgtgatgctcgtcaaggggggcgagcctatgga 5350
|||||
```

```
Db 6052 gccacctcgaactgagcgtcgatctttgtgatgctcgtagggggcgagcctatgga 6111
|||||
QY 5351 aaacgccagcaacgcgcctttttacgggttctcgtggttctggttctggtcctttgctcaca 5410
|||||
Db 6112 aaacgccagcaacgcgcctttttacgggttctcgtggttctggtccttttgcctcaca 6171
|||||
QY 5411 tgtttctcgtggttatccctgattcctgtgtgataacgattatccgctttagtgag 5470
|||||
Db 6172 tgtttctcgtggttatccctgattcctgtgtgataacgattatccgctttagtgag 6231
|||||
QY 5471 ctgataccgctcgcgcagccgaacacgcagcagcagcagcagcagcagcagcagcagc 5530
|||||
Db 6232 ctgataccgctcgcgcagccgaacacgcagcagcagcagcagcagcagcagcagcagc 6291
|||||
QY 5531 aagagcgcccaatacagcaaacccctctcccgcggttgccgattcattaatgcagct 5590
|||||
Db 6292 aagagcgcccaatacagcaaacccctctcccgcggttgccgattcattaatgcagct 6351
|||||
QY 5591 ggcaacgacaggtttcccgactggaaaagcgggcagtgagcgcaacgcaattaatgtgatt 5650
|||||
Db 6352 ggcaacgacaggtttcccgactggaaaagcgggcagtgagcgcaacgcaattaatgtgatt 6411
|||||
QY 5651 agctcaactaattagcaccacccaggtttacacatttatgcttccggctcgtatgtgtg 5710
|||||
Db 6412 agctcaactaattagcaccacccaggtttacacatttatgcttccggctcgtatgtgtg 6471
|||||
QY 5711 gaattgtgagcgatacaaatccacaggaataacagcagctatgacctgattacgccca 5767
|||||
Db 6472 gaattgtgagcgatacaaatccacaggaataacagcagctatgacctgattacgccca 6528
|||||

RESULT 8
AAV02043/C
ID AAV02043 standard; DNA; 6295 BP.
XX
AC AAV02043;
XX
DT 08-JUN-1998 (first entry)
XX
DE Plasmid pWRG3196 encoding murine Interleukin-12.
XX
KW Interleukin-12; IL-12; cytokine; growth factor; mouse
KW Plasmid pWRG3196; cancer; tumour; metastasis; gene therapy; ds;
KW cyclic; circular.
XX
OS Chimeric - Mus musculus.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Bos taurus.
OS Chimeric - Rhesus macaque polyoma virus
OS Chimeric - encephalomyocarditis virus.
XX
FH Key Location/Qualifiers
CDS 955..1675
FT /*tag= a
FT /product= p35 subunit
FT /note= "contains an Intron"
FT Intron 1261..1333
FT /*tag= b
FT CDS 2377..3384
FT /*tag= c
FT /product= p40 subunit
XX
PN WO9746263-A1.
XX
PD 11-DEC-1997.
XX
PF 04-JUN-1997; 97WO-US09591.
XX
PR 05-JUN-1996; 96US-0659206.
XX
PA (AURA-) AURAGEN INC.
XX
```

PI Rakhmillevich AL, Yang N;
XX WPI: 1998-041898/04.
DR P-PSDB; AAW44004-05.
XX
PT Interleukin-12 gene therapy of tumours - comprises delivering
PT construct containing promoter and sequences encoding interleukin-12
PT p35 and p40 subunit(s) to target cells in vivo
XX
PS Claim 8; Page 33-38; 50pp; English.
XX
CC Plasmid pWRG3196 is a bicistronic plasmid encoding both subunits,
CC i.e. p35 (see AAW44004) and p40 (see AAW44005), of murine interleukin-12
CC (IL-12). The p35 and p40 subunits were cloned from a mouse spleen
CC cDNA library. The vector contains a single cytomegalovirus promoter,
CC an SV40 splicing donor/splicing acceptor, and bovine growth hormone
CC polyA signal. Between the p35 and p40 genes is an internal
CC ribosome entry site element (IRES) cloned from encephalomyocarditis
CC virus. The backbone of the plasmid is from pUC19. Plasmid
CC pWRG3139 induces about half the expression of IL-12 as the tandem
CC vector pWRG3169 (see AAV02042) in vivo and in vitro. A novel method
CC of treating tumours in a mammal involves delivering copies of an
CC pWRG3196, comprising a promoter operative in the mammalian
CC epidermal cells and DNA sequences encoding p35 and p40 subunits of
CC IL-12 to target cells in vivo. Delivery of the construct allows
CC IL-12 expression for treatment of solid, metastatic or disseminated
CC tumours, and regression of established tumours. The treatment is
CC effective even when the genetic construct is delivered to a site
CC distant from the tumour.
XX
SQ Sequence 6295 BP; 1610 A; 1610 C; 1568 G; 1507 T; 0 other;

Query Match 45.6%; Score 2630; DB 19; Length 6295;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3138 actgagcgcgtgtttatcaaacgtctgactgggaacacctggcgttaccacctaactg 3197
DB 6295 ACTGCCCTGCTGTTTACAACTGCTGACTGGGAAACCTGGCGTTACCCAACTTAATCG 6236
QY 3198 ccttcagcacatccctcttcgcagctggcgtaataagcgaagagcccgaccgacgcg 3257
DB 6235 CATTGCACACATCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCGCCGATCG 6176
QY 3258 ccttcaccaacagttgcgcagcctgaatggcggaatggcgctgatgcggtattttctcct 3317
DB 6175 CCTTCCCAACAGTTGCGCAGCCTGAATGGGGAATGGCGCCTGATGCGGTATTTCTCCT 6116
QY 3318 tacgcatctgtggtgtatttacacccgcacatattgtgcactctcagtaacaatctgtctga 3377
DB 6115 TACGGATGTGCGGTATTTACACCCGATATGTTGTCACCTCTCAGTACAACTGCTCTGA 6056
QY 3378 tgcgcagtagtaagccagcccgacacccgcgaacaccccgctgacgcgacctgacgagc 3437
DB 6055 TGCCGCATAGTTAAGCCAGCCCGACACCCGCCAACCCCGCTGACGCCCTGACGGGC 5996
QY 3438 ttgtctgtcccggtacccgtcttacagacaagctgtgacgctctccggagagctgctgtg 3497
DB 5995 TTGTTCTGCTCCGGCATCCGTTACAGACAGCTGTGACCGTCTCCGGGAGCTGATGTG 5936
QY 3498 taagagggttttaccgctcatcaccgaaacgcgcgcgagacgaaggcctcgtgatacgct 3557
DB 5935 TCAGAGGTTTTCACCGTCTATCACCAGAAACGGCGAGAGAAAGGCCCTCGTATACGCT 5876
QY 3558 atttttataggttaatgcatgataataatggtttcttagcgtcaggtggcaactttcgc 3617
DB 5875 ATTTTATAGGTTAATGTCATGATTAATGTTTCTTAGACGTGAGGTGGCAGCTTTTCG 5816
QY 3618 gggaatgtgcgcggaacccctattgtttatttttctaaaaacattcaaatatgatctcc 3677
DB 5815 GGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTTAATATACATTCANATATGTATCC 5756

QY 3678 gctcatgagacaataaacctgataaaatgcttcaataataattgaaaaaggaagagtatgag 3737
DB 5755 GCTCATGAGACAATAAACCTGATAAAATGCTTCAATAAATATTGAAAAGGAAGAGTATGAG 5696
QY 3738 tattcaacattccgtgtcgcccttatcccttttttctggcgcatatttcccttccctgtttt 3797
DB 5695 TATTCAACATTCCTGTCGCCCTTATTTCCCTTTTGGGGGCAATTTTGGCTTCCCTGTTT 5636
QY 3798 tgcaccccaaaaacgctggtgaaagtaaaagatgctgaagatcgtggtgggtgcacgagt 3857
DB 5635 TGCTCACCCAGAAACGCTGGTGAAGATGAAGATGCTGAAGATCAGTTGGGTGCACGAGT 5576
QY 3858 ggggttaactgaacttgatctcaacagcggtaagatccttggagagtttctgcgccgaaga 3917
DB 5575 GGGTTACATCGAACCTGGATCTCAACAGCGGTAAAGTCTTGAGAGTTTTCGCCGCCAAGA 5516
QY 3918 acgttttccaatlgatgagcaccttttaaagtctctatgtggcggtatttatcccgat 3977
DB 5515 AGCTTTTCCATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATATATCCCGTAT 5456
QY 3978 tgaacccggggaagagcaactcgggtcgccgcatacacatacttctcagaatgacttggttga 4037
DB 5455 TGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGA 5396
QY 4038 gtactcaccagtcacagaaaaagcatcttaccggatggcagatgacagtaagagaattatgag 4097
DB 5395 GTACTCACCACTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTTGCGAG 5336
QY 4098 tgcctccataaccatgagtataaacactgcccgaacttactctcgaacacgacgtcgagg 4157
DB 5335 TGCTGCCATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5276
QY 4158 accgaagagctaacccgcttttttgcacaacatgggggagatcatgaaactcgcttgatcg 4217
DB 5275 ACCGAAGGAGCTAACCGCTTTTTCGACAAACATGGGGATCATGTAACCTGCGCTTGATCG 5216
QY 4218 ttgggaacccggagctgaaatgaagccataccaaaacagcagcgtgacacacacagatcctgt 4277
DB 5215 TTGGGAACCCGGAGCTGAATGAAGCCATACCAAACAGCAGCGTGACACCAAGTCCCTGT 5156
QY 4278 agcaatggcaacaacgctgtgcgaactattaaactggcgaacttacttacttacttctcccg 4337
DB 5155 AGCAATGGCAACAGCTTGGCGCAAACTATTAACTGGCGNACTACTTACTTCTAGTCTCCG 5096
QY 4338 gcaacaattaatagactggatggagcgagataaagtgtgcagaccacttctcgctcgcg 4397
DB 5095 GCAACAATTAATAGACTGGATGGAGCGCGATAAAGTTGCAGGACCACCTTCTGCGCTCGGC 5036
QY 4398 ccttcggcgtggctggtttattgtctgataaaatctggagccggtgagcgtgggtctcgcgg 4457
DB 5035 CTTTCCGCTGGCTGGTTTATTTGCTGATAAATCTGGAGCCGCTGAGCGTGGGTCTCGCG 4976
QY 4458 tatcattgcagcactggggccagatggtaagccctcccgatcgtagttatctacacgac 4517
DB 4975 TATCATTCGACGACCTTGGGGCCAGATGGTAAGCCCTCCCGTATCTAGTATCTACACGAC 4916
QY 4518 ggggggtcagggaacactatggatgaacgaataagacagatcgctgagataggtggcctcact 4577
DB 4915 GGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT 4856
QY 4578 gattaagcattggtaactgtcagacccaagtttactcataatactatttagattgattaaa 4637
DB 4855 GATTAAAGCATTTGGTAACGTGCAGACCAAGTTTACTCATATATATTTAGATTGATTAAA 4796
QY 4638 acttcatttttaatttaaaggatcaggtgagatccttttataatctctcatgacaca 4697
DB 4795 ACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCTCTTTTGTGATATCTCATGACCAA 4736
QY 4698 aatcccttaacgtgagtttcttcactgagcgtcagaccccgtagaagaagatcaag 4757
DB 4735 AATCCCTTAACGTGAGTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAANAAGATCAAGG 4676

QY 4758 atcttcttgagatccctttttttctgcgcgtaatactgctgcttgcgcaaaaaaacacc 4817
DB 4675 ATCTTCTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGCACAAAACACAC 4616
QY 4818 gctaccagcgggtgttcttgcgcgatacaagagctaccaactcttttccgaagtaac 4877
DB 4615 GCTACAGCGGTGTGTTGTTGCGGATCAAGAGCTACCAACTCTTTTCCGAAGTAAC 4556
QY 4878 tggcttcagcagcagcagacataccaaatactgtccttctagtgtagcgttagtgcca 4937
DB 4555 TGGCTTTCAGCAGCGCAGATACCAAAATCTGTCTTCTAGTGTAGCGGTAGTGGCCA 4496
QY 4938 caacttcaagaactctgttagcagccgctacatacctcgtctgctctgcttaactctgttaccagt 4997
DB 4495 CCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTTCGCTCTGCTATCTGTACAGT 4436
QY 4998 ggcctgctgcagtgccgataaagtcgtcttaccgggttggaactcaagcagatagtacc 5057
DB 4435 GGTCTGCTGCAGTGGCGATAAGTCGTCTTACC GG GTTGGA CTCAAGACGATAGTTACC 4376
QY 5058 ggataaggcgcagcggctcgggtgaaacgggggttcgtgcacacagccagcttggagcg 5117
DB 4375 GGATAAGCGCAGCGCTCGGGCTGTAACGGGGGTTCGTGCACACACCCAGCTTGGAGCG 4316
QY 5118 aacgaactcacaccgaactgagatacctacagcgtgagctatgagaagcgcacgcttcc 5177
DB 4315 AACGACTACACCGAACTGAGATACCTACAGCTGAGCTATGAGAAGGCCAGCTTCC 4256
QY 5178 cgaaggagaaagcgcagcagctatccgtgaagcgcagggcgggaacagagagagcgcac 5237
DB 4255 CGNAGGAGANAGCGGACAGGTATCCGTAGCGGAGGTGCGAAGAGAGAGCGCAC 4196
QY 5238 gggggagcttcacgggggaaacgcctgggtatctttagctgctgctgggttgcacact 5297
DB 4195 GAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTTAGTCTGCTGCGGTTCGCCACCT 4136
QY 5298 ctgacttgagcgtcgatcttctgagtcgctcagggggggcggagcctatgaaaaaacgc 5357
DB 4135 CTGACTTGAGCTCGATTTTGTGATGCTCGTCAAGGGGGCGGAGCTATGAAAAACGC 4076
QY 5358 cagcaacgcggcctttttagcgtctcctgcttctttagcgtcttctgctcactgtctt 5417
DB 4075 CAGCAACGGCGCTTTTACGGTCTCTGGCCCTTTGCTGGCCCTTTGCTCAATGTTCTT 4016
QY 5418 tctgcttatccctgattctgttgataaccgctattaccgcttttgagtgagctgatac 5477
DB 4015 TCTGCTTATCCCTGATCTGTGATAACCGTATTACCGCTTTGAGTGAGCTCATAC 3956
QY 5478 cgtcgccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5537
DB 3955 CGCTCGCCGACGCCGAACAGCGAGCGCAGCAGTCACTGAGCGAGGAAGCGAAGCG 3896
QY 5538 cccaatcagcaaacgcctctccgcgcgttggcgcgttcaatgagcgtgagcagcagcagc 5597
DB 3895 CCAATACGAACCGCTCTCCCGCGGTGGCGGATTCATTATGAGCTGGCACCA 3836
QY 5598 cagggttcccgactggaagcggcagtgagcgcgaacgaacgaacgaacgaacgaacgaacga 5657
DB 3835 CAGGTTTCCGACTGGAAGCGGCGAGTGAAGCGCAACGAATTAATGTGAGTTAGCTCAC 3776
QY 5658 tcattaggacccccagcgtttacactttatgctccgctgctgctgctgctgctgctgctg 5717
DB 3775 TCATTAGGACCCCGGCTTACATTTATGCTTCGGCTCGCTATGCTTGTGGAATTGT 3716
QY 5718 ggcgcgatacaatttcacacaggaacagcgtatgacctgattacgcga 5767
DB 3715 GAGCGGATACAAATTCACACAGGAACAGCTATGACCAATGATTACGCCA 3666

RESULT 9
AAD10238/c
ID AAD10238 standard; DNA; 7164 BP.
XX

AC AAD10238;
XX 24-SEP-2001 (first entry)
XX Commercial plasmid vector pCMVbeta.
XX Plasmid; cyclic; circular; replicon; exogenous gene; marker gene;
KW transcription termination; immunostimulatory sequence; ISS; antiviral;
KW non-essential nucleotide; molecular biology application; gene therapy;
KW DNA vaccine; cloning; gene expression; in vitro protein production;
KW cytosstatic; pCMVbeta; cytomegalovirus promoter; lacZ gene;
KW ampicillin gene; simian virus 40; SV40 intron; ds.
XX Chimeric - Cytomegalovirus.
OS Chimeric - Rhesus macaque polyoma virus;
OS Chimeric - Unidentified.
XX Key Location/Qualifiers
FT misc_feature 1..26
FT /*tag= a
FT /note= "Corresponds to the non-essential nucleotide
FT sequence that have been removed in the novel DNA
FT plasmid vector"
FT 4513..4890
FT misc_feature /*tag= b
FT /note= "pUC18 lacZ promoter that corresponds to the
FT non-essential nucleotide sequence that have been removed
FT in the novel DNA plasmid vector"
FT 7007..7164
FT /*tag= c
FT /note= "5' non-coding region of lacZ gene from pUC18 that
FT corresponds to the non-essential nucleotide sequence
FT that have been removed in the novel DNA plasmid vector"
XX WO200151626-A2.
XX 19-JUL-2001.
XX 09-JAN-2001; 2001WO-US01255.
XX 10-JAN-2000; 2000US-0480879.
XX (ELIM-) ELIM BIOPHARMACEUTICALS INC.
XX Lu X, Sun L, Zhang Y;
XX WPI; 2001-451855/48.
XX New plasmid DNA vectors, useful for most molecular biology
PT applications, e.g. gene therapy, DNA vaccines, cloning and expression
PT of genes, and in the in vitro production of polypeptides and/or
PT proteins
XX Example 2; Page 42-44; 50pp; English.
XX The present invention relates to plasmid DNA vectors comprising
CC essentially of a replicon and at least one other component selected from
CC promoter, intron, exogenous gene, transcription termination sequence,
CC selectable marker gene, detectable marker gene and an immunostimulatory
CC sequence (ISS), where the non-essential nucleotide sequences have been
CC substantially removed from these vectors. The plasmid DNA vectors are
CC useful in most molecular biology applications, e.g. gene therapy, DNA
CC vaccines, cloning and expression of genes, and in the in vitro production
CC of polypeptides and/or proteins. The present sequence is a commercial
CC plasmid DNA vector pCMVbeta which comprises Cytomegalovirus promoter,
CC Simian Virus 40 (SV40) intron, the lacZ gene, pUC origin of replication
CC and the ampicillin gene.
SQ Sequence 7164 BP; 1700 A; 1840 C; 1897 G; 1727 T; 0 other;

Query Match 45.6%; Score 2630; DB 22; Length 7164;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	3138	actggcgcgtctttacaacgctgactgagtggaagaaacccctgcggtaccacaactaatcgc	3197	
Db	7164	ACTGGCGTCTGTTTACAACTGCTGACTGGGAAACCTTGGCGTTACCCAACTTAATCG	7105	
Qy	3198	ccttgacgacacatcccctttccagctgcggttaataagcgaagagcccgacccgacgc	3257	
Db	7104	CCTTGCAGCACAATCCCTTTTGGCAGCTGGCGTAATAGCGAAGAGGCCCGACCGATCG	7045	
Qy	3258	cccttcccaacagttgcgcagcctgaaaggcgaatggcgccctgtagcggtattttctcct	3317	
Db	7044	CCCTTCCCAACAGTTGGCAGCCCTGAATGGGGAATGGCGCCTGATGCGGTATTTCCT	6985	
Qy	3318	tacgcatctgtcggttatttcaaccgcgaatatggtgaactctcagtaacaatctcctga	3377	
Db	6984	TACGCATCTGTGCGGTATTTCACACCGCATATGGTGACACTCTCAGTACAATCTGCTCTGA	6925	
Qy	3378	tgcgcatagttaagccagcccgacacccgcgaaccccgctgacgcgcctgacgggc	3437	
Db	6924	TGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGGCCCTGACGGGC	6865	
Qy	3438	ttgtctgcctccggcactcgtctacagacaagctgtgaccgtctccggagctgcattg	3497	
Db	6864	TTGTCGTCTCCGGCATCCGCTTACAGACAAGCTGTGACCCGTCTCCGGGAGCTGCATGTG	6805	
Qy	3498	tcagaggttttcacgcgctacccgaaacgcgcgagacgaaggccctcgatagcgcct	3557	
Db	6804	TCAGAGGTTTTTCACCGCTCATCCGGAACCGCGGAGAGAGGAGGCGCTCGGTGATACGCT	6745	
Qy	3558	atttttagttaaagtcataataatggtttcttagcgtcaggtgcgcaacttttcg	3617	
Db	6744	ATTTTATAGGTTAATGTATGTAATAATGGTTCTTAGACGTCAGGTGCACTTTTCG	6685	
Qy	3618	gggaaatgctgcggaaacccctattgtttattttctaaatacattcaaatatgtatcc	3677	
Db	6684	GGGAAATGTGCGCGGAACCCCTATTGTTTATTTCCTAATACATTCAAATATGTAATCC	6625	
Qy	3678	gctcatgagacaataaccctgataaaatgcttcaataatattgaaaggaagagatagag	3737	
Db	6624	GCTCATGAGACAAATACCCCTGATAATGCTTCAATAATATTGAAAGGAAGAGATATGAG	6565	
Qy	3738	tattcaacatttcgctgcgccttattcccttttttgcgcattttgccttcctgtttt	3797	
Db	6564	TATTCAACATTTTCGCTGTCGCCCTTATTCCCTTTTTCGCGCATTTTGCCCTTCCTGTTT	6505	
Qy	3798	tgctcacccagaacgcctggtgaagtaaaagatgctgaagatcagttgggtgcacgagt	3857	
Db	6504	TGCTCACCCAGAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT	6445	
Qy	3858	gggttacatcgaactgatctcaacagcggtaagatccttgagagtttcgccccgaaga	3917	
Db	6444	GGGTTACATCGAATGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCGGAAGA	6385	
Qy	3918	acgttttccaatgatgacacatttaagttctgctatgtgcgcggtattattcccgat	3977	
Db	6384	ACGTTTCCAATGATGACACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTAT	6325	
Qy	3978	tgcgcggggcaagagacaactcgtgcgcgcatacacattattctcagaatgacttggtga	4037	
Db	6324	TCACCGCGGCAAGAGCAACTCGTTCGCCGCATACACTATTCTCAGAAATGACTTGGTTGA	6265	
Qy	4038	gtactcaccagtcacagaaaagcatcttaacgtaggcgatgagtaagagaaatattgcag	4097	
Db	6264	GTACTCACCAGTCAAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTAATGAG	6205	
Qy	4098	tgctgcataaccatgagtataacactgcgcgcgaacttacttctgacaacgatcggagg	4157	
Db	6204	TGCTGCCATATACCATGATGATACACTTGCGGCCCAACTTACTTCTGACAAAGATCGGGAGG	6145	
Qy	4158	accgaagagctaaccgccttttttgacaacatgggggatactgtaactcgccttgatcgc	4217	
Db	6144	ACCGAAGAGGTAACCCGCTTTTTCACAAACATGGGGGATCATGTAACCTGCCCTTGATCG	6085	

Qy	4218	ttgggaacccggagctgaatgaagccaataccaaacgacgagcgtgacacacgagtgcctgt	4277	
Db	6084	TTGGGAACCGGAGCTGAATGAAGCATACCAAAACGACGAGCGTGACACCGATGCCTGT	6025	
Qy	4278	agcaatggcaacaacgctgcgcaaaactattaactggcggaactactactactagcttcgcg	4337	
Db	6024	AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTCTAGCTTCCCG	5965	
Qy	4338	gcaacaataatagactggatggagggcggaataaagttagagaccacacttcgcgctcgcg	4397	
Db	5964	GCAACAATTAATAGACTGGATGGAGGGCGGATAAAGTTCCAGAGCACACTTTCGGCTCGGC	5905	
Qy	4398	ccttcgcgctgcctgggtttatttgctgataaatactggagccggtgagcgtgggtctcgcg	4457	
Db	5904	CCTTCGCGCTGCCTGGTATTATTTGCTGATAAATCTGAGCGGTGAGCGTGGGTCTCGCGG	5845	
Qy	4458	tatcattgcacacactggggccagatggaagccctcccgatctcgtagttatctacacgac	4517	
Db	5844	TATCATTTGCACACTGGGGCCAGATGTTAAGCCCTCCCGTATCGTAGTATTCTACACGAC	5785	
Qy	4518	ggggagtcaggcaactatggatgaacgaaaatagacagatcgtcagataggtgcctcact	4577	
Db	5784	GGGGAGTCAGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT	5725	
Qy	4578	gattaagcattggttaactgtcagacccaagtttactcattatactattagatttataa	4637	
Db	5724	GATTAAAGCATTTGTAACCTGTGACAGCAAGTTTACTCATATATACTTTAGATTGATTAA	5665	
Qy	4638	acttcatttttaatttaaaagatcctagtggaagatccttttatactatcctagaccaa	4697	
Db	5664	ACTTCATTTTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCATGACCAA	5605	
Qy	4698	aatcccttaacgctgagtttctccactgagcgtcagaccccgtagaaaaagatacaagg	4757	
Db	5604	AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCACACCCCGTAGAAAAAGATCAAAGG	5545	
Qy	4758	atctcttgagatccttttttctgcggttaactctgctgtgcacacacacacacacaccc	4817	
Db	5544	ATCTCTTTGAGATCCTTTTCTCGCGTAACTGCTGTGTCGTAACAAAAAACCACCACC	5485	
Qy	4818	gctaccagcgtggtttgtttgcggatcaagagctcaaacactcttttcgaaggtaac	4877	
Db	5484	GCTACCAGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGGAAGGTAAAC	5425	
Qy	4878	tggtctcagcagcgcagataaccaaaactgtcctcttagtagcgttagtaggcca	4937	
Db	5424	TGGCTTCAGCAGAGCGGCAGATACCAAAATCTGCTTCTAGTAGCCGTAGTTAGGCCA	5365	
Qy	4938	ccaactcaagaactctgttagccgcctacatacctcgtcctcgttaactcctgttaccagt	4997	
Db	5364	CCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTTAATCTCTTACCAGT	5305	
Qy	4998	ggctgctccagtggtggagataagtgcttaccgggttggaactcaagacagatagttacc	5057	
Db	5304	GGCTGCTCCCAAGTGCGGATAAGTCTGCTTACCGGGTTGGACTCAAGACCATAGTTTACC	5245	
Qy	5058	ggataagcgcgcagcgttcgggctgaacgggggttcgtgcacacagcccgcttgagcgc	5117	
Db	5244	GGATAAGCGCGAGCGGTGCGGCTGAACGGGGGTTCGTGTCACACAGCCAGCTTGGAGCG	5185	
Qy	5118	aacgacctacaccgaactgagatacctacagcgtgagcta tgaaaaagcgcgaagcttcc	5177	
Db	5184	AACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCACGCTTCC	5125	
Qy	5178	cgaggagagaagcgcgacaggtatccggttaagcgcgaggtcggaacagagagcgcac	5237	
Db	5124	CGAAGGGAGAAGCGGACAGGTATCCGGTAAGCGGAGGGTCGGAAGAGAGAGCGCAC	5065	
Qy	5238	gaggagacttccaggggaaacgcctggtatctttatagctcgtcggggttcgcacact	5297	
Db	5064	CAGGAGCTTCCAGGGGAAACGCGCTGTTATCTTTTATAGTCTCTGCGGTTCGCCACCT	5005	

Db 2275 GCCCAATACGCAACCGCTCTCCCGCGCTTGGCGGATTCATTATGACAGCTGGCACG 2216
 Qy 5597 acaggtttcccgactgaaagcggcagtgagcgcaacgcaatgaattagtgtagctca 5656
 Db 2215 ACAGGTTTCCGACTGGAAGACGGCGAGTAGAGCAACGCAATTAATGTGAGTTAGCTCA 2156
 Qy 5657 ctcattaggaaccccgagctttacacittatgcttcgcgctcgtatgtgtggaattg 5716
 Db 2155 CTCATTAGGCAACCCAGGCTTTACACTTATGCTTCGCGCTCGTATGTGTGGAATTG 2096
 Qy 5717 tgagcgataacaatttcacacaggaacagctatgacatgattacg 5764
 Db 2095 TGAGCGGATACATTTTCACACAGGAACAGCTATGACCATGATTACG 2048

RESULT 13

AAA95416

ID AAA95416 standard; DNA; 5733 BP.

XX AAA95416;

XX 12-FEB-2001 (first entry)

XX Aspergillus nidulans pAN52-1 vector.

DE Fusions protein; aqueous two-phase system; ATPS; protein production;
 KW protein purification; expression library screening; vector; ds.
 XX Aspergillus nidulans.

XX key Location/Qualifiers

FT promoter 1..2129

FT /tag= a

FT CDS 2130..2304

FT /tag= b

FT /product= "gpdA"

FT terminator 2305..3071

FT /tag= c

XX WO200058342-A1.

XX 05-OCT-2000.

XX 24-MAR-2000; 2000WO-FI00249.

XX 25-MAR-1999; 99FI-0000667.

XX 20-AUG-1999; 99FI-0001782.

XX (VALM) VALTION TEKNIILINEN TUTKIMUSKESKUS.

XX Penttilae M, Nakari-Setaelae T, Fagerstroem R, Selber K, Kula M;
 PI Linder M, Tjernerfeld F;
 XX WPI; 2000-686858/67.

XX Isolation and purification of proteins or cells in aqueous two-phase
 PT systems, comprises combining a desired protein or a cell to a targeting
 PT protein capable of isolating the protein or cell into one of the phases

XX Example 1; Page 93-95; 109pp; English.

XX The present invention is related to a novel method for separating,
 CC purifying and isolating proteins and cells. This involves the use of
 CC liquid-liquid extraction in an aqueous two-phase system (ATPS) which
 CC partitions molecules by fusing them to targeting proteins which then
 CC carry the molecules of interest into one of the phases. The present
 CC sequence is an expression vector which was used to demonstrate the method
 CC of the invention. The method is useful in also useful in the
 CC identification of nucleic acid sequences in expression library screening.

XX Sequence 5733 BP; 1435 A; 1454 C; 1378 G; 1463 T; 3 other;

Query Match 45.6%; Score 2628; DB 21; Length 5733;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3137 cactggcgctgcttttacacgctgactgggaaacccctggcttacccaacttaac 3196
 Db 3095 cactggcgctgcttttacacgctgactgggaaacccctggcttacccaacttaac 3154
 Qy 3197 gccttgagcacatcccccctttcgccagctggcgtaagagagagcccgaccgac 3256
 Db 3155 gccttgagcacatcccccctttcgccagctggcgtaagagagagcccgaccgac 3214
 Qy 3257 gccttcccaacagttgcgagcctgaaatggcgaaatggcgctgatgcggtattttctcc 3316
 Db 3215 gccttcccaacagttgcgagcctgaaatggcgaaatggcgctgatgcggtattttctcc 3274
 Qy 3317 ttacgcatctgctgggtattttcacacccgcataatggtgcactctcagtaacaatctgcttg 3376
 Db 3275 ttacgcatctgctgggtattttcacacccgcataatggtgcactctcagtaacaatctgcttg 3334
 Qy 3377 atgcgcatagttaagccagccccgcacacccgcgaaccccgctgacgcgccctgacggg 3436
 Db 3335 atgcgcatagttaagccagccccgcacaccccgctgacgcgccctgacggg 3394
 Qy 3437 ctgtctgctcccgcatccgctttacagacaagctgtgacgctcccgaggagctgcatgt 3496
 Db 3395 ctgtctgctcccgcatccgctttacagacaagctgtgacgctcccgaggagctgcatgt 3454
 Qy 3497 gtcagaggttttcacccgtcatcccgaaacgcgcgagacgaaaggccctgtgatacgc 3556
 Db 3455 gtcagaggttttcacccgtcatcccgaaacgcgcgagacgaaaggccctgtgatacgc 3514
 Qy 3557 tattttatagtttaatgctatgataataatggtttctctagacgttcaggctgaccttttc 3616
 Db 3515 tattttatagtttaatgctatgataataatggtttctctagacgttcaggctgaccttttc 3574
 Qy 3617 ggggaaatgtgcgaggaaacccctattgtttatttttataatacatcacaatattgtatc 3676
 Db 3575 ggggaaatgtgcgaggaaacccctattgtttatttttataatacatcacaatattgtatc 3634
 Qy 3677 cgctcatgagacaataaccctgataataatgcttcaataataattgaaaaagggaagatga 3736
 Db 3635 cgctcatgagacaataaccctgataataatgcttcaataataattgaaaaagggaagatga 3694
 Qy 3737 gtattcaacatttcgctgctccctattcccttttttgcggcattttgcctctcctgtt 3796
 Db 3695 gtattcaacatttcgctgctccctattcccttttttgcggcattttgcctctcctgtt 3754
 Qy 3797 ttgctcaccggagaaacgcctggtgaaagttaaagatgctgaagatcagttgggtgcacgag 3856
 Db 3755 ttgctcaccggagaaacgcctggtgaaagttaaagatgctgaagatcagttgggtgcacgag 3814
 Qy 3857 tgggttaacatgaactggatctcaacagcggtaagatccttgagaggttttcgcccgaag 3916
 Db 3815 tgggttaacatgaactggatctcaacagcggtaagatccttgagaggttttcgcccgaag 3874
 Qy 3917 aacgttttcccaatgatgacgacttttaagttctgctatgtggcgcggtattattcccgta 3976
 Db 3875 aacgttttcccaatgatgacgacttttaagttctgctatgtggcgcggtattattcccgta 3934
 Qy 3977 ttgacgcgggcaagagacaaactcggtccgcgcatacatatttctcagaatgacttggttg 4036
 Db 3935 ttgacgcgggcaagagacaaactcggtccgcgcatacatatttctcagaatgacttggttg 3994
 Qy 4037 agtaactaccagtcacagaaaagcatcttacggtatggcatgacagtaagagaattatgca 4096
 Db 3995 agtaactaccagtcacagaaaagcatcttacggtatggcatgacagtaagagaattatgca 4054
 Qy 4097 gtctgcgcataaccatgagtataacactgcggccaacttacttctgacaacgacgag 4156
 Db 4055 gtctgcgcataaccatgagtataacactgcggccaacttacttctgacaacgacgag 4114

PS Example 2; Page 79-89; 123pp; English.

XX The invention describes novel genes and vectors exhibiting increased
 CC expression and novel splicing patterns. It provides a gene encoding a
 CC Factor VIII protein, that comprises one or more consensus or near
 CC consensus splice sites which have been corrected to increase expression.
 CC The method, DNA sequences and expression vectors can be used to increase
 CC the expression of a gene, especially a factor VIII gene. Genes containing
 CC modified 5' and/or 3' untranslated regions have optimized expression
 CC levels and tissue-specific expression. The methods are used for
 CC identification and correction of consensus splice sites, addition of
 CC introns, optimization of 5' and 3' untranslated regions and increase in
 CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
 CC to treat a clinical disorder, to study RNA processing and/or gene
 CC regulation. The present sequence represents an uncorrected version of
 CC beta-domain deleted factor VIII protein encoding gene sequence (construct
 CC pc9-2). This was used to develop a new coding sequence for beta-domain
 CC deleted factor VIII protein by correcting the consensus splice sites.
 XX

SQ Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;

Query Match 45.6%; Score 2628; DB 20; Length 9164;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3137 cactggcgctgttttacaacgtgctgactgggaaacccctggcgttacccaactaatc 3196
 Db 9147 CACTGGCGCTGTTTACACAGCTGCTGACTGGGAAACCCCTGGCGTTACCCNACTTAATC 9088
 Qy 3197 gcttgacgacatccccctttccagctgctgtaataagcgaagagcccgaccgac 3256
 Db 9087 GCTTGGCAGACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCGCACCGATC 9028
 Qy 3257 gcccttcccaacagtgtgcgacgctgaatggcggaatggcgctgactgctgtttctcc 3316
 Db 9027 GCCCTTCCCAACAGTGGCGACCTGAATGGCGAATGGCGCTGATGGGTATTTCTCC 8968
 Qy 3317 ttacgactgtgctgggtatttacaacgcgcatatgtgtcactctcagtaacatctgctgtg 3376
 Db 8967 TTACGCACTGTGCGGTATTTACACCGCGATATGCTGCACTCTCAGTACATCTGCTCTG 8908
 Qy 3377 atgcgcatagttaagccagcccgacaccccgcaacaccccgctgacgcccctgacggg 3436
 Db 8907 ATGCCGATAGTTAAGCCAGCCCGCACCCCGCCCAACACCCGCTGACGCGCCCTGACGGG 8848
 Qy 3437 ctgtctgtccccggcatccgcttacagacaagctgtgacgctctccgggagctgcatgt 3496
 Db 8847 CTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGT 8788
 Qy 3497 gtcagaggttttcccgctcatcaccgaaacgcgcgagacgaaaggccctgtgatacgcc 3556
 Db 8787 GTCAGAGGTTTTCACCGCTCATCACCGAAGCCGCGAGACGAAAGGGGCTCGTGATAGGCC 8728
 Qy 3557 tattttataggttaagtcatgataataatgtttcttagacgtcaggtggcacttttc 3616
 Db 8727 TATTTTATAGTTAATGTCATGATAATATGTTTCTTAGACGTCAGGTGGCACTTTTC 8668
 Qy 3617 ggggaatatgtgcgcgaacccctattgtttattttcttaatacatctcaaatatgtatc 3676
 Db 8667 GGGGAATATGTCGCGGAACCCCTATTGTTTATTTTCTTAATACATTCAAATATGTATC 8608
 Qy 3677 cgtcatgagacataaccctgataaagtcttcaataatattgaaaaaggagagatga 3736
 Db 8607 CGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGAGTATGA 8548
 Qy 3737 gtattcaacatttcgctgctcccttattcccttttttggcgacttttgccttctctgttt 3796
 Db 8547 GTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTCGGGCATTTTGGCCTTCTCTGTTT 8488
 Qy 3797 ttgcttccccgagaacgcgtgtgaaagtataaagatgctgaaagatcagttgggtgcacag 3856
 Db 8487 TTGCTCACCCAGAAACCGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 8428

Qy 3857 tgggttacatcgaaactggatctcaacagcggtgaagatcccttgagagttttcgccccgaag 3916
 Db 8427 TGGTTCATCGAACTGGATCTCAACAGCGGTAAAGTCTTGGAGAGTTTTCGCCCGGAAG 8368
 Qy 3917 aacgttttccaatgatgagacactttttaaagttctgctatgtggcgcggtatatcccgta 3976
 Db 8367 AACGTTTTCGAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGGCTATATATCCCGTA 8308
 Qy 3977 ttgacgcggcggaagagcaactcggtcgccgcatcacactattctcagaatgacttggctg 4036
 Db 8307 TTGACGCGGCGCAGACCACTCGGTCCGCCGATACACTATTCTCAGANAATGACTTGGTTG 8248
 Qy 4037 agtactcaccagtcacagaaaaagcatcttaagtgatggaatgacagtaagaattatgca 4096
 Db 8247 AGTACTCACCAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATATGCA 8188
 Qy 4097 gtgctgcataaacatgagtgataaacactcgcgccaacttactctgacaacgactcgag 4156
 Db 8187 GTGCTGCCATAAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAACGATCGGAG 8128
 Qy 4157 gaccgaaggagtaacccgctttttgcacaacatgggggagcatgtaactcgcccttgatc 4216
 Db 8127 GACCGAAGGAGCTAACCCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCCTTGATC 8068
 Qy 4217 gttgggaacccggagctgaatgaagccataccaaacgcgagcgtgacacccagatgcctg 4276
 Db 8067 GTTGGGAACCCGGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACCCAGATGCCTG 8008
 Qy 4277 tagcaatggcaaacagcttgcgcaaacatttaactggcgaaactacttacttagcttccc 4336
 Db 8007 TAGCAATGGCAACAGCTTGGCAAACTATTAACTGGCGAACTACTTACTACTAGCTTCCC 7948
 Qy 4337 ggcaacaatgaactgagctggatgagcgaggaataaagtgcagagaccactctcgctcgg 4396
 Db 7947 GGCACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGAGACCCTTCTGGCTCGG 7888
 Qy 4397 cccctccggctggctgtttatctgataaactctgagccggtgagcgtgggtctcgcg 4456
 Db 7887 CCTTCCGGCTGGCTGTTTATGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCGCG 7828
 Qy 4457 gtacattgacgactgggcccagatggaagccctccgtactctgtagttatctacacga 4516
 Db 7827 GTATCTTGCAGCACTGGGCCAGATGTAAGCCCTCCCGTATCTGATCTATCTACACGA 7768
 Qy 4517 cggggagtcagcgaactatgagtaacgaaatagacagatcgctgagatgagtgccctcac 4576
 Db 7767 CGGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTGCCTCAC 7708
 Qy 4577 tgattaagcatgtgtaactgtcagaccaaagtttactcataataatacttttagattttaa 4636
 Db 7707 TGATTAAGCATTTGTAAGTGTGAGACCAAGTTTACTCATATATATCTTTAGATTGATTTAA 7648
 Qy 4637 aacttcatttttaatttaaaagatcagtggaagatcccttttgaataatctcatgacca 4696
 Db 7647 AACTTCATTTTTTAATTAATAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATGACCA 7588
 Qy 4697 aaatcccttaacgtgagtttctccactgagcgtcagaccccgtagaagaatcctaaag 4756
 Db 7587 AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAG 7528
 Qy 4757 gatcttttgagatactcttttttctgctgctgctgctgctgctgctgctgctgctgctg 4816
 Db 7527 GATCTTCTGAGATCCTTTTTTCTGCGGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7468
 Qy 4817 cgttaccagcgt 4876
 Db 7467 GCGTACCAGCGGT 7408
 Qy 4877 ctgcttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 4936
 Db 7407 CTGCTTCAGCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7348

QY 3437 cttgtctctcccgcatccgcttacagagaagctgtgacccgtctcccgaggagctcatgt 3496
DB 9330 cttgtctctcccgcatccgcttacagagaagctgtgacccgtctcccgaggagctcatgt 9271
QY 3497 gtcagaggtttccacgctcatccacgaaacgcgcgagacgaaagggcctcgtgatacgc 3556
DB 9270 gtcagaggtttccacgctcatccacgaaacgcgcgagacgaaagggcctcgtgatacgc 9211
QY 3557 tatttttaggttaattgcatgataataatggtttctcttagacgtcaggtggcacttttc 3616
DB 9210 tatttttaggttaattgcatgataataatggtttctcttagacgtcaggtggcacttttc 9151
QY 3617 ggggaaatgctgcgggaaacccctatttatttttataataataataataataataata 3676
DB 9150 ggggaaatgctgcgggaaacccctatttatttttataataataataataataataata 9091
QY 3677 cgtcatgagacaataaaccttgataaaatgcttcaataatattgaaaaaggaagataga 3736
DB 9090 cgtcatgagacaataaaccttgataaaatgcttcaataatattgaaaaaggaagataga 9031
QY 3737 gtattcaacatttcgctgcgccttattcccttttttggggcattttgaccttctgttt 3796
DB 9030 gtattcaacatttcgctgcgccttattcccttttttggggcattttgaccttctgttt 8971
QY 3797 ttgctcaacccagaaacgctgtgaaagttaaaagatgctgaagatcagttgggtgcacgag 3856
DB 8970 ttgctcaacccagaaacgctgtgaaagttaaaagatgctgaagatcagttgggtgcacgag 8911
QY 3857 tgggttacctggaactgatactcaacagcgttaagatccttgagagtttttcgccccgaag 3916
DB 8910 tgggttacctggaactgatactcaacagcgttaagatccttgagagtttttcgccccgaag 8851
QY 3917 aacgttttccaatgtagagcaacttttaagttctctgtagtggcgggtattatcccgta 3976
DB 8850 aacgttttccaatgtagagcaacttttaagttctctgtagtggcgggtattatcccgta 8791
QY 3977 ttgagccggggaagagaactcgtgcccgcatacacattcttcagaaatgacttggttg 4036
DB 8790 ttgagccgggggaagagaactcgtgcccgcatacacattcttcagaaatgacttggttg 8731
QY 4037 agtactcaccagctcacagaaagcattctacggatggatggatgacagtgaagaaattatga 4096
DB 8730 agtactcaccagctcacagaaagcattctacggatggatggatgacagtgaagaaattatga 8671
QY 4097 gtgctgccataaccatgtagtaaacactgcggccaacttacttctgacaacgactcgag 4156
DB 8670 gtgctgccataaccatgtagtaaacactgcggccaacttacttctgacaacgactcgag 8611
QY 4157 gaccgaaggagctaacccgttttttgacaacatgggggatacatgtaactgccttgatc 4216
DB 8610 gaccgaaggagctaacccgttttttgacaacatgggggatacatgtaactgccttgatc 8551
QY 4217 gttggaaacccgagctgaatgaagccataccaaacgagagcgtacaccacgactgcctg 4276
DB 8550 gttggaaacccgagctgaatgaagccataccaaacgagagcgtacaccacgactgcctg 8491
QY 4277 tagcaatggcaacaacgtttgcgaacactattaaactggggaacttacttactagcttccc 4336
DB 8490 tagcaatggcaacaacgtttgcgaacactattaaactggggaacttacttactagcttccc 8431
QY 4337 ggcacaattaatagactggaatggggcggaataaagtgtgaggaacaacttctgcgtcgg 4396
DB 8430 ggcacaattaatagactggaatggggcggaataaagtgtgaggaacaacttctgcgtcgg 8371
QY 4397 ccttcccgctgctgggtttattgtgataaaatcggagccgtgagcgtgggtctcgag 4456
DB 8370 ccttcccgctgctgggtttattgtgataaaatcggagccgtgagcgtgggtctcgag 8311
QY 4457 gtaatttcagcactggggccagatggtaagccctcccgtatcgtagttatctacaga 4516
DB 8310 gtaatttcagcactggggccagatggtaagccctcccgtatcgtagttatctacaga 8251

QY 4517 cgggaggtcaggcaactatggatgaacgaataagacagatcgctgagatagtgacctcac 4576
DB 8250 cgggaggtcaggcaactatggatgaacgaataagacagatcgctgagatagtgacctcac 8191
QY 4577 tgattaaagcattggttaactgtcagaccagtcttactcatataacttttagattgatttaa 4636
DB 8190 tgattaaagcattggttaactgtcagaccagtcttactcatataacttttagattgatttaa 8131
QY 4637 aacttcatttttaatttaaaagatctagggtgaagatcctttttgataatctcatgacca 4696
DB 8130 aacttcatttttaatttaaaagatctagggtgaagatcctttttgataatctcatgacca 8071
QY 4697 aaatcccttaacgtgagtttctccactgagcgtcagacccttagaanaagatacaag 4756
DB 8070 aaatcccttaacgtgagtttctccactgagcgtcagacccttagaanaagatacaag 8011
QY 4757 gatctcttgagatcctttttctgcgcgtaactctgcgtctgcaaaaaaaccacac 4816
DB 8010 gatctcttgagatcctttttctgcgcgtaactctgcgtctgcaaaaaaaccacac 7951
QY 4817 cgtaccagcgggtggtttgtttgcccggatcaagagctaccaactcttttccggaagtaa 4876
DB 7950 cgtaccagcgggtggtttgtttgcccggatcaagagctaccaactcttttccggaagtaa 7891
QY 4877 ctggcttcaagagagcgcagataccaaaatactgctcctctagtgtagccgttagtgcc 4936
DB 7890 ctggcttcaagagagcgcagataccaaaatactgctcctctagtgtagccgttagtgcc 7831
QY 4937 accactcaagaaactctgtagcaccgcctacatacctcgtctgcttaactcctgtaccag 4996
DB 7830 accactcaagaaactctgtagcaccgcctacatacctcgtctgcttaactcctgtaccag 7771
QY 4997 tggctgctccagtgccgataaagtgcgtcttaccgggttggaactcaagacgatagttac 5056
DB 7770 tggctgctccagtgccgataaagtgcgtcttaccgggttggaactcaagacgatagttac 7711
QY 5057 cggataaaggcgcagcggctcgggctgaacgggggttcgtgcacacagccagcttgagc 5116
DB 7710 cggataaaggcgcagcggctcgggctgaacgggggttcgtgcacacagccagcttgagc 7651
QY 5117 gaacgacctacacccgaactgagatacctacagcgtgagctatgagaaagccacgcttc 5176
DB 7650 gaacgacctacacccgaactgagatacctacagcgtgagctatgagaaagccacgcttc 7591
QY 5177 ccgaagggaaggaagcggagcaggtatccggtaagcggcgggttcggaacagagagcgca 5236
DB 7590 ccgaagggaaggaagcggagcaggtatccggtaagcggcgggttcggaacagagagcgca 7531
QY 5237 cagaggagcttccaggggaaacgcctggtatcttatagtcctgcgggttccgcaacc 5296
DB 7530 cagaggagcttccaggggaaacgcctggtatcttatagtcctgcgggttccgcaacc 7471
QY 5297 tctgactgagcgtcgatcttctacgggttcctggccttttgcggccttttgcacatgtct 5356
DB 7470 tctgactgagcgtcgatcttctacgggttcctggccttttgcggccttttgcacatgtct 7411
QY 5357 ccagcaaacgcggccttttctacgggttcctggccttttgcggccttttgcacatgtct 5416
DB 7410 ccagcaaacgcggccttttctacgggttcctggccttttgcggccttttgcacatgtct 7351
QY 5417 ttctcgttataccctgattctgtgataacacgtattaccgccttttgagtgagctgata 5476
DB 7350 ttctcgttataccctgattctgtgataacacgtattaccgccttttgagtgagctgata 7291
QY 5477 ccgctcgcgcagcgaacacgagcagcagcagcagcagcagcagcagcagcagcagcagc 5536
DB 7290 ccgctcgcgcagcgaacacgagcagcagcagcagcagcagcagcagcagcagcagcagc 7231
QY 5537 gcccaatacgaacacgcctctcccgcgcttggtggcgaattcattatgcagctggcagc 5596
DB 7230 gcccaatacgaacacgcctctcccgcgcttggtggcgaattcattatgcagctggcagc 7171
QY 5597 acaggtttcccactgggaagcgggagtgagcgcgaacgcaattaatgtgagttagctca 5656

```
|||||
Db 7170 ACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTCTCA 7111
Qy 5657 ctcataggcacccagcctttacacttttatgcttcggcgtcgatgtgtgtggaatg 5716
Db 7110 CTCATTAGGCACCCCGAGGCTTTACACTTTTATGCTTCCGGCTCGTATGTTGTGGGAATTG 7051
Qy 5717 tgagcggataacaatttcacaggaacacagctatgaccatgattacg 5764
Db 7050 TGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACG 7003
```

Search completed: August 31, 2002, 22:58:22
Job time: 16803 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 19:58:51 ; Search time 6064.54 Seconds
(without alignments)
12834.764 Million cell updates/sec

Title: US-09-810-861b-3
Perfect score: 5767
Sequence: 1 agctgcattgcctgcaggtc.....ctatgaccatgattacgccca 5767

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size: 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	18.4	2080	11 BC001541	BC001541 Homo sapi
2	864	15.0	1070	9 AJ281552	AJ281552 4A3A-P6F1
3	841	14.6	841	9 AL042026	AL042026 DKFZp434E
4	835	14.5	1067	9 AU081137	AU081137 AU081137
5	828	14.4	1089	9 AU081124	AU081124 AU081124
6	723	12.5	954	9 AL044364	AL044364 DKFZp434C
7	723	12.5	1004	9 AJ281480	AJ281480 4A3A-P4G8
8	691	12.0	1013	10 BM438846	BM438846 IPLVr0015
9	691	12.0	1039	9 AU081040	AU081040 AU081040
10	686	11.9	780	10 BI753192	BI753192 603026066
11	682	11.8	1163	9 AU081044	AU081044 AU081044
12	679	11.8	680	9 AL646751	AL646751 AU081044
13	678	11.8	707	9 AL656688	AL656688 AU081044
14	678	11.8	800	9 AJ281449	AJ281449 4A3A-P4D5
15	673	11.7	724	9 AL645114	AL645114 AU081044
16	672	11.7	715	9 AL661706	AL661706 AU081044
17	662	11.5	705	9 AL643164	AL643164 AU081044

C 18	659	11.4	659	9	AL643220
C 19	659	11.4	689	9	AL646532
C 20	657	11.4	847	12	AZ687169
C 21	652	11.3	653	9	AL662063
C 22	648	11.2	648	9	AL640650
C 23	648	11.2	703	9	AJ281437
C 24	647	11.2	767	9	AL040542
C 25	645	11.2	645	9	AL639182
C 26	645	11.2	670	9	AL662130
C 27	645	11.2	717	12	AZ208376
C 28	642	11.1	642	9	AL642207
C 29	642	11.1	670	9	AL635952
C 30	640	11.1	640	12	B84895
C 31	638	11.1	832	10	BG923768
C 32	636	11.0	687	9	AU001481
C 33	633	11.0	688	9	AL042640
C 34	629	10.9	657	9	AL641508
C 35	628	10.9	669	9	AL660789
C 36	627	10.9	705	9	AL635845
C 37	624	10.8	637	9	AL627526
C 38	623	10.8	675	9	AL636713
C 39	621	10.8	647	9	AL642844
C 40	619	10.7	645	9	AL642835
C 41	619	10.7	655	9	AL659614
C 42	619	10.7	664	9	AL639797
C 43	617	10.7	650	9	AL638248
C 44	613	10.6	672	9	AL640799
C 45	612	10.6	663	9	AU001472

ALIGNMENTS

RESULT 1

BC001541
LOCUS BC001541 2080 bp mRNA linear HTC 31-JAN-2002
DEFINITION Homo sapiens, clone IMAGE:3453362, mRNA.
ACCESSION BC001541
VERSION BC001541.1 GI:14705895
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2080)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdedpaxil.stanford.edu
R. M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, D. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 4 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156
This clone has the following problem: no polyA-tail.
Location/Qualifiers
1. 2080
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:3453362"
/tissue_type="Placenta, chorioncarcinoma"
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 319 a 692 c 651 g 418 t
ORIGIN

Query Match 18.4%; Score 1064; DB 11; Length 2080;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 844 ccccgagtgctgtgcacagccttcctggcttccccctcttcctcctcctcctc 903
Db 104 CCCCCGAGTGTGTGTGCACAGCCTTCCCTGGCTTCCCACACTCTTCTCTCCTCCTC 163
QY 904 tggctctctgggtggagagtgagggtgagggccgggagatgcagagtgctgtgacg 963
Db 164 TGGCTCTGGGTGGAGAGTGGGGGTGAGGCCGGGAGGATGCAGAGCTGTGTGTGAGC 223
QY 964 gtgctggggccgctcgaggcattcctcctgaagaccgccggggccctgtctctgt 1023
Db 224 GTGGTGGGGCCGGCTGCGGGGCATTCGCCCTGAAGACCCCGGGGGCCCTGTCTGTCT 283
QY 1024 ttccctgggcatccctttgggagccaccatgggacccctgcctttctgcacccggag 1083
Db 284 TTCTCTGGGATCCCTTTCCGGGAGCACCATGGGACCCCGTCTGCTTCTGCCACCGGAG 343
QY 1084 cccaagcagccttggtcaggggtggtagacgtacacaccttcacagtgctgtctaccaa 1143
Db 344 CCAAAGCAGCCTTGTGACGGGTGTAGACGCTTACAACCTTCCAGAGTGTCTGTACCAA 403
QY 1144 tatgtgacacacctatcacaggttttgaggccaccgacagatgtggaaccccaacgtgag 1203
Db 404 TATGTGGACACCTTATCCACAGGTTTGTAGGACCGGAGATGTGGAAACCCCAACCGTGA 463
QY 1204 ctgagcgagagctgctgttacctcaagctgtggaacaccatccccggcctacatcccc 1263
Db 464 CTGAGCGAGAGCTGCTGTACCTCAACGTGTGACACCATATCCCCGGGCTACATCCCCC 523
QY 1264 accctgtcctcgtctggaatctatgggggtggtcttctacagtggggcccctcctctggag 1323
Db 524 ACCCTGTCTCTGTCTGTGATCTATGGGGGTGCTTCTACAGTGGGGCCCTCCTCTGGAC 583
QY 1324 gtgtacgatggcgctctctgtacagccgagagagactgtgtgtgtccatgaactac 1383
Db 584 GTGTACGATGGCGGCTCTTGTGTACAGCCGAGAGGAGCTGTGTGTGTCCATGAACACTAC 643
QY 1384 cgggtgggagcctttggtctctgcttgccttgcctggggagccgagagcccgggcaatgtg 1443
Db 644 CGGGTGGGAGCCTTTGGCTTCTTGGCCCTTCCGGGAGCCGAGAGCCCGGGCAATGTG 703
QY 1444 ggtctcctgatacagaggtggtcctcgtcagtgagtgagagaaacgtgagccttcggg 1503
Db 704 GGTCTCTCTGGATCAGAGGCTGGCCCTGTGAGTGGGTGTGAGGAAACGTGCGACCCCTCGGG 763
QY 1504 ggtgacccgacatcagtgacgtgtgttgggagagcggggagccctcgtgggcatg 1563
Db 764 GTGTACCCGACATCAGTGACGTGTGTGGGAGAGCGCGGGAGCCCTCGGTGGGCATG 823
QY 1564 caccctgtctccccccagccggggcctgttccacagggcccgctgtgcagagcgtgtgcc 1623
Db 824 CACCTGTCTCCCCGCCAGCCGGGGCCCTTCCACAGGGCCGTGTGTCAGAGCGGTGCC 883
QY 1624 cccaatggacctgggccaacgttgggcatgggagagggcccgctcagggcccaacgtgtg 1683
Db 884 CCAATGGACCTGTGGCCACAGTGGGCAATGGGAGAGGGCCCGCTCGCAGGGCCACGACGTG 943
QY 1684 gccacacctgtgggtgtctcctcagggcactgtgtgggaatgacacagagctgtagcc 1743
Db 944 GCCACCTGTGGGTGTCTCTCCAGCGGCACCTGTGTGGGAATGACACAGAGCTGTGTAGCC 1003
```

```
QY 1744 tgcttcggacacgacagcagcgaggtccctgggtgaaccacgaatggcacgtgctgctcaa 1803
Db 1004 TGCTTCGACGACGACGACGAGGTCTTGGTGAACCAAGATGGACGCTGCTGCTCAA 1063
QY 1804 gaaagcgttctccggttctcctctcgtgcctgtggtgtagatggagacttctcagtgacac 1863
Db 1064 GAAAGCGTCTTCGGTCTCTCTTCGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1123
QY 1864 ccagagggcctcatcaacgaggagagacttccagcgctcaggt 1907
Db 1124 CCAGAGGCCCTCATCAACGCGGAGACTTCCAGGCGCTGCAGGT 1167
```

RESULT 2
AJ281552/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

1..1070

/organism="Anopheles gambiae"

/strain="4A r/z"

/db_xref="taxon:7165"

/clone="4A3A-P6F11"

/clone_lib="Anopheles gambiae immune competent 4A3A"

/cell_line="immune competent 4A3A"

/lab_host="E. coli DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from

forward priming site which reads from the 3' end of the

cDNA. The 4A3A is a directionally cloned and normalized

cDNA library that was constructed from the 4A3A cell line

oligo-T primed cDNA according to: Bonaldo, Lennon & Soares

(1996) : Normalization and Subtraction: Two approaches to

Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 263 a 283 c 255 g 269 t

ORIGIN

Query Match 15.0%; Score 864; DB 9; Length 1070;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 964; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4398 ccttcggcgtcgtcgtttattgctgataaaatctggagccggtgagcgtggtctcggg 4457

Db 966 CTTCCGGCTGCTGCTGCTTTATTGCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCGCG 907

QY 4458 tatcattgcagcactggggccagatggttaagccctcccgctatcgtagttatctacacgac 4517

Db 906 TATCATTTGACGACATGGGGCCAGATGGTAAAGCCCTCCCGTATCGTAGTTTATTCTACACGAC 847

Qy	4518	gggagtcaggcaactatgagatgaacgaataagacagatcgctgagatagtgctcaact	4577
Db	846	GGGAGTCAGGCAACTATGATGGAGCAANAATAGACAGATCGCTGAGATAGTGCTCACT	787
Qy	4578	gattaacattggttaactgtcagaccgaagttaactcatatactattagattgattaaa	4637
Db	786	GATTAAAGCATTTGTTAACTGTGACAGCAAGTTTACTCATATATACTTTAGATTGATTAAA	727
Qy	4638	acttcatttttaatttaaaaggatctagggtgaagatcctttttgataatctcatgaccaa	4697
Db	726	ACTTCATTTTAAATTTAAAGAGATCTAGTGAAGATCCTTTTGTATATCTCATGACCAA	667
Qy	4698	aatcccttaacgtgagtttctgctccactgagctcagacccttagaagaagatcaaaag	4757
Db	666	ANTCCCTTAACGTGAGTGTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAGG	607
Qy	4758	atctcttgagatccttttttctgcgcgtaactgtcgtctgtctgcaacaaacaaacacc	4817
Db	606	ATCTTCTTGAGATCCTTTTCTTCGCGGTAATCTGCTGTGCAACAAACAAACACCACC	547
Qy	4818	gtcaccagcgggtgtttgtttgcccagatcaagagctaccaactcttttccgaaggtaac	4877
Db	546	GCTACCAGCGGTGTGTGTTTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAAAGGTAAC	487
Qy	4878	tggcttcagcagcgcagataccataactctctctctctgtagcgttagtgagcca	4937
Db	486	TGGCTTCAGCAGAGCGGAGATACCAATACACTGTTCTTCTAGTGTAGCGGTAGTAGGCCA	427
Qy	4938	ccacttaagaactctgtagcaccgcctacatacctcgtcgtctgtctaaacctgtttaccagt	4997
Db	426	CCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCTCTGTCTAATCTGTTACCACT	367
Qy	4998	ggctgtcgcagtgggcgataagtcgtctcttaccgggttggaactcaagacgatagttacc	5057
Db	366	GGCTGCTGCCAGTGGCGATAGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACC	307
Qy	5058	ggataaagcgcagcgtggcgtgaacgggggttcgtgcacagccagccttggagcg	5117
Db	306	GGATAAGCGCGCGGTGCGGCTGAACGGGGGTTGCTGACACAGCCGACCTTTGGAGCG	247
Qy	5118	aacgacctacaccgaactgagataccctacagcgtgagctatgagaagcgcacgccttcc	5177
Db	246	AACGACCTACACCGAAGCTGAGATACCTACACGGTGAGCTATGAGAAAGCGCCAGCTTCC	187
Qy	5178	cgaggggagaagcgcagcaggtatccggtgaagcggcgaggttcggaaacagagagcgcac	5237
Db	186	CGAAGGGAGAAGCGGACAGGTATCGGTAAGCGGCGAGGGTTCGGAACAGGAGAGCGCAC	127
Qy	5238	gaggagcttccagggggaacgcctggtatctttatagtcctcgtcggttcccaact	5297
Db	126	GAGGAGCTTCCAGGGGGAACGCTGTATCTTTATAGTCTCTGTCGGGTTTCGCCACCT	67
Qy	5298	ctgacttgagcgtcgatttttctgctcgtcaggggggagggcgagcctatggaaaaacgc	5357
Db	66	CTGACTTGAGCGTCGATTTTGTGATGCTGCTCAGGGGGGCGGAGCCCTATGGAAGAACGC	7
Qy	5358	cagcaa 5363	
Db	6	CAGCAA 1	
RESULT 3			
AL042026			
LOCUS			
DEFINITION			
AL042026			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

REFERENCE

1 (bases 1 to 841)

Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

EST (Poustka, et al.)

Unpublished (1999)

Contact: Poustka A.J.

Department Lebrach

Max-Planck-Institute for Molecular Genetics

Innestrasse 73, 14195 Berlin, Germany

Tel: +49-30-84131623

Fax: +49-30-84131128

Email: poustka@mpg-berlin-dahlem.mpg.de

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp434E111) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1. 841

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKFZp434E111"

/clone_lib="434 (synonym: htes3)"

/tissue_type="testis"

/dev_stage="adult"

/lab_host="DH10B"

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT

226 a 192 c 202 g 221 t

ORIGIN

Query Match

14.6%; Score 841; DB 9; Length 841;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

3639

tatttggttatttttctaataacattcaaatatgtatccgctcatgagacataaaccttg

3698

Db

1

TATTGTTTATTTCCTAAATACATTCAAATATGATCGCTCATGACACATAACCTTG

60

QY

3699

ataaatgcttcaataatattgaataagagagatgtagtattcaacatttccgtgtgc

3758

Db

61

ATAAATGCTTCAATAATATTGAAAAGGAGAGATGATGATTTCAACATTTCCGTGCGC

120

QY

3759

ccttattccctttttgcgcattttgcttccctgtttttgtcaccacagaaacgtggt

3818

Db

121

CCTTATTCCTTTTTCGGCATTTTGCTTCTCTCTTTCCTCACCAGAACCTGCT

180

QY

3819

gaagtaaaagatgctgaagatcagttgggtgcagcagtggttaccatcgaaactgacct

3878

Db

181

GAAAGTAAAGATGCTGAAGATCAGTTGGTGCACAGCTGGGTTACATCGAACTGGATCT

240

QY

3879

caacagcggtaagatccttgagagtttccgcccgaagaacgctttcccaatgatgagcac

3938

Db

241

CAACAGCGGTGAAGATCCTTGAGAGTTTTCGCCCGCAAGAACGTTTTCCTCAATGATGAGCAC

300

QY

3939

ttttaaaagtctgctatgtgcgcggtattatccggtattacgcggttcacgcgagagcaact

3998

Db

301

TTTTAAAGTCTGCTATGTGGCGCGGTATATCCCGTATTGACGCGGGCAAGACACT

360

QY

3999

cggctgcgcgcatacactattctcagaatgacttggttgtagtactcaccagtcacagaaaa

4058

Db

361

CGGTGCGCGCATACACTATTCAGAAATGACTTGGTTGAGTACTCACCAGTCACAGAAA

420

QY

4059

gcattctacgagtgagcagtaagagaattatgcagtgctgccaataaccatgagtgca

4118

Db

421

GCATCTTACGGATGGCATGACAGTAAGAGATTTATGCAAGTGTGCGCATTAACCATGAGTGA

480

QY

4119

taacactcggccaacttacttctgacaacgatcgaggagaccgaagagcgttaacccgttt

4178

```
Db 481 TAACACTGCGGCCAACTTACTTCTGACAAAGATCGGAGGACCGAAGGAGCTAACCCGCTTT 540
QY 4179 ttgcaacaatgggggacatgaactgccttgatcgttggaaccggagctgaatga 4238
Db 541 TTGTCACAATGGGGGATCATGTAACCTGCCTTGATCGTTGGGAACCGAGCTGAATGA 600
QY 4239 agccatacaaacagcagcgcgtgacaccacgatgcctgtagcaaatggcaaacagcttgcg 4298
Db 601 AGCCATPACCAAAACGAGCGGTGACACACGATGCCGTAGCAATGGCAACACGTTGCG 660
QY 4299 caaactattaaactggcgaactactactacttccgtccgcgaacaattaatagactggaat 4358
Db 661 CAAACTATTAACTGGCGAACTACTTACTAGCTTCCCGCAACAATTAATAGACTGGAT 720
QY 4359 ggaggcggataaagtgcaggaccacttgcgctgcgcctcccgctggctggctggttat 4418
Db 721 GGAGCGGATAAAGTTGCAAGGACCACTTCTGCGCTCGGCCCTTCGCGCTGGCTGTTAT 780
QY 4419 tgcgtataatctggcgcggcgagcgtggtggttcctgcgtatcattgcagcactggggcc 4478
Db 781 TGCATGATAAATCTGGAGCGCGTGAGCGTGGTCTCGCGGTATCATTTGCAGCACTGGGGCC 840
QY 4479 a 4479
Db 841 A 841

RESULT 4
AU081137/c 1067 bp mRNA linear EST 30-AUG-2001
LOCUS AU081137 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncoirhynchus mykiss cDNA clone K12,
mRNA sequence.
ACCESSION AU081137
VERSION AU081137.1 GI:6431485
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncoirhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncoirhynchus.
1 (bases 1 to 1067)
Kono,T., Sakai,M. and LaPatra,S.E.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
Contact: Wasahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
Location/Qualifiers
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="K12"
/tissue="Kidney"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"

BASE COUNT 235 a 292 c 288 g 250 t 2 others
ORIGIN

Query Match 14.5%; Score 835; DB 9; Length 1067;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 935; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4831 gtttggccggatcaagagctaccacactctttccgaagtaactggcttcagcaga 4890
Db 1020 GTTTGTTTCCCGATCAACAGCTACCAACTCTTTTCCGAAGTAAGTGGCTTCAGCACA 961
```

```
QY 4891 ggcagataccacaaatactgtctcttagttagcccttagttagggccaccacttcaagaac 4950
Db 960 GGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGTGTAGCCACCACTTCAAGAAC 901
QY 4951 tctgtagcaccgcctacatacctcgtctcgttaactcctgttaccagtggtcgtgcagt 5010
Db 900 TCTGTAGCACCGCCTACATACCTTCGCTCTGCTTAATCCTGTTACCACTGGCTGCTGCCAGT 841
QY 5011 ggcgataaagtctgtcttaccgggttggactcaagcagatagttaccgagataaaggcgag 5070
Db 840 GCGGATAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTACCGGTAAGGCGCAG 781
QY 5071 cggctcgggtgaacgggggttcgtgcacacagccagcttggagcgaacgacctacac 5130
Db 780 CGGTCTGGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTGGAGGAACGACCTACACC 721
QY 5131 gaactgagatacctacagcgtgagctatagaaaaagcgcaagccttcccgaaggagaaaag 5190
Db 720 GAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAGCTTCCGGAAGGAGAGAAAG 661
QY 5191 ggcgacaggtatcccggttaagcggcagggctcggaacagagagcgacgagaggettcca 5250
Db 660 GCGGACAGGTATCCGCTAAGCGGCGAGGTCGGAACAGGAGGCGCACGAGGGAGCTTCCA 601
QY 5251 gggggaaacgcctggtatcttttatagtcgtggtgttgcacacctctgacttgagct 5310
Db 600 GGGGAAACGCGCTGTTATCTTTATAGTCTGCGGTTTCGCCACCTCTGACTTGAGCGT 541
QY 5311 cgaattttgtgatcgtcagggggcgagcctatgaaaaaacgcaacgagcgccg 5370
Db 540 CGATTTTGTGATGCTGTCAGGGGGCGAGGCTATGGAABACGCCAGCAACACGCGGCC 481
QY 5371 ttttacggttccctggccttttgcgtggttttgcacatgttcttctcgcttatcc 5430
Db 480 TTTTACGCTTCTCGCCCTTTTGTGCGCTTTTGTCTCACATGTTCTTCTCGGTTATCC 421
QY 5431 cctgattctgtggataaccgattaccgctttagtgagctgataccgctgcgcgacg 5490
Db 420 CCTGATCTCTGGATACCGCTATTACCGCTTTGAGTGAGCTGATACCGCTCGCCGACG 361
QY 5491 cgaacacccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5550
Db 360 CGAACACCCAGCGCAGCGAGTCAGTGAGGAGAAAGCGGAGCGCCCAATACGCAAA 301
QY 5551 ccgctctctcccgcgctgttgccgattcattaatgagctggaagcagcagcagcagcagc 5610
Db 300 CGGCTCTCCCGCGGCTTGGCGGATTCATTATATGACGCTGGCAGCAGAGTTTCCCGAC 241
QY 5611 tggaaagcggcagtgagcgcaacgcaataatgtgagtagctcactcattagcacc 5670
Db 240 TGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGCTACTCATTTAGGCACC 181
QY 5671 caggctttacatttatgcttcggctcgtatgtgtgtggaattgtgagcgaatacaa 5730
Db 180 CAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAAATGTGAGCGGATAACAA 121
QY 5731 ttccacaggaacagctatgacctgattacgcca 5767
Db 120 TTTCACACAGGAACAGCTATGACCATGATTAGGCCA 84

RESULT 5
AU081124/c
LOCUS AU081124 Oncoirhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncoirhynchus mykiss cDNA clone KG'12,
mRNA sequence.
ACCESSION AU081124
VERSION AU081124.1 GI:6431472
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncoirhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Actinopterygii; Neopterygii; Teleostei; Euteleostomi;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 1089)
Kono,T., Sakai,M. and LaPatra,S.E.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.

Location/Qualifiers
1. .1089
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KG12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"
245 a 297 c 295 g 251 t 1 others

BASE COUNT
ORIGIN

Query Match 14.4%; Score 828; DB 9; Length 1089;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4940 acttcaagaactctgtagcaccgcctacatacctcgtctgtcgttaacgagcagatagttaccgg 4999
Db 912 ACTTCAGAACTCTGACACCGCTACATACCTCGCTCTGCTAATCTGTTACCAAGTGG 853
Qy 5000 ctgttcgagtgccgataagtcgtgtcttaccgggttggaactcaagacagatagttaccgg 5059
Db 852 CTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTGGACTCAAGACGATGTTACCGG 793
Qy 5060 ataagggcagcggctcgggctgaacggggggttcgtgcacacagcccgcttgagcgaa 5119
Db 792 ATAAGGCGCAGCGGTGGGGCTGAACGGGGGGTTCGTGCACACAGCCCGCTTGGAGCGAA 733
Qy 5120 cgacctacacgaactgagatcctacacgctgagcgtgagctatgagaaagcccgcttcgg 5179
Db 732 CGACCTACACGAAGTACCTACAGCTGAGCTATGAGAAAGCCCGCTTCCCG 673
Qy 5180 aaggagaaaaggcgacagggtatccggtaagcggcagggctgggaacagagagcgacga 5239
Db 672 AAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGTCGGAACAGAGAGCGCACGA 613
Qy 5240 gggagcttccagggggaaacgcctggtatctttagtctcgtcggttccgcacctct 5299
Db 612 GGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGTCTGTCCGGTTTCGCCACCTCT 553
Qy 5300 gacttgagctgcgatttttggatcgtcgtcagggggcgagcctatggaaaacgcga 5359
Db 552 GACTGACGCTCGATTTTGTGTGATCTCGTCAGGGGGCGGAGCTATGGAANAACGCCA 493
Qy 5360 gaacgcggccctttttacgggttcctggcccttttgcgtcccttttgcctcacatgtcttcc 5419
Db 492 GCAACGGCGCCTTTTACGGGTTCTTGGCCTTTTGTGGCCTTTTGTCTCACATGTTCTTTC 433
Qy 5420 ctgcggtatccccgtatctctgtgataaccgtattacacgcctttgagtgagctgatacgg 5479
Db 432 CTGCGTTTATCCCTGATTCTGTGGATACCTGATTTACCGCCTTTTTCAGTGAGCTGATACCG 373
Qy 5480 ctccgcgcagccgaacccagcgacgagtcagtgagcagagagaaagcgaagcgc 5539
Db 372 CTCGCCGCGCGGACAGACCGGACCGAGTCTAGTGAGCGAGGAAGCGGACGCC 313
Qy 5540 caatacgcnaaacccgcctctccccgcgcttggccgattcattaaagcagctggcgacga 5599
Db 312 CAATACGCAAAACCGGCTCTCCCGCGCGCTTGGCCGATTTCATTATGACAGCTGGCAGACA 253

Qy 5600 ggtttccgcgactgaaagcgggcaagtgagcgcaacgaataatgagtagtcactc 5659
Db 252 GGTttccgcgactgaaagcgggcaagtgagcgcaacgaataatgagtagtcactc 193
Qy 5660 attagccaccagcgtcttacactttatgctctccgctcgatgctgtgtggaattgga 5719
Db 192 ATTAGGCACCCAGGCTTTACACTTTATGCTTCCCGCTCGCTATGTGTGGAATTGGA 133
Qy 5720 gcggataacaatttcacacagagaaacagctatgaccatgattacgcca 5767
Db 132 GCGGATAACAATTTTCACACAGAGAAACAGCTATGACCATGATTACGCCA 85

RESULT 6
AL044364/c
LOCUS
DEFINITION DKEZp434C172_s1 434 (synonym: htes3) Homo sapiens CDNA clone
DKEZp434C172_3', mRNA sequence.
ACCESSION AL044364
VERSION AL044364.1 GI:5432586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 954)
Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Ansorge, Benes, et al.)
Unpublished (1999)
Contact: Ansorge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKEZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. .954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKEZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSport1; Site_1: NotI; Site_2: SalI"
214 a 273 c 253 g 214 t

BASE COUNT
ORIGIN

Query Match 12.5%; Score 723; DB 9; Length 954;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 773; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4815 accgctaccagcggctggtttgttgcgggatacaagagctaccacactctttccgaagt 4874
Db 954 ACCGCTACCAGCGGTGTTGTTGTCGGGATCAAGAGCTACCAACTCTTTTCCGAAGGT 895
Qy 4875 aactggttcagcagagcgagatcaccaataactctctctagtgtagccgtagtagg 4934
Db 894 AACTGGCTTCAGCAGCGCGAGATACCAAAATACTCTTCTTAGTGTAGCCGTAGTAGG 835
Qy 4935 ccaccacttcaagaactctgttagcggcgcctacactcgcctctgcttaactcgttacc 4994
Db 834 CCACCACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCTCTGCTAANTCCTGTTACC 775

```
QY 4995 agtggctgctgcagtgccgataaagtcgtgtcttaccgggttggaactcaagacgatgtt 5054
Db 774 AGTGGCTGCTGCCAGTGGCGGATAAAGTCGTCTCTACCGGGTTGGACTCAAGACGATAGTT 715
QY 5055 accggataagcgccagcggtcgccggctgaacgggggttcgtgcacacagcccgcttggga 5114
Db 714 ACCGGATAAGCGCCAGCGGTCGGGCTGAACGGGGGTTGCTGCACACAGCCCAAGCTTGGGA 655
QY 5115 gcgaacgacctacacgaactgagatcactacagcgtgagctgagctatgagaagcccgct 5174
Db 654 GCGAAGCGACCTACACCGAAGTACCTACAGCGTGAGCTATGAGAAAGCCCGACGCT 595
QY 5175 tcccgaaggagaaagcggacaggttatccggttaagcggcaggggtcggaaacagagagcg 5234
Db 594 TCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGCTCGGAACAGAGAGCG 535
QY 5235 caagagggagctccaggggggaaacccctggtatctttagtccctgctcgggttccgca 5294
Db 534 CAGAGGGAGCTTCCAGGGGGAACCCCTGGTATCTTTATAGTCTCTCGGGTTTCGCCA 475
QY 5295 cctctgacttgagcgtcgatcttctgagctcgtcagggggcgagccctatggaaaaa 5354
Db 474 CTTCTGACTTGACGTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAAAAA 415
QY 5355 cgcgaacagcgccgtttttttagcgttcctggtccttttgcgtccttttgcctacatgtt 5414
Db 414 CGCCAGCAACGGCGGCTTTTACGGTTCCTGGGCTTTTGTGCGCTTTTGTGCACATGTT 355
QY 5415 cttctcggttatccctgattctgtgataaccgtattaccgctttgagtgagctga 5474
Db 354 CTTTCTCGGCTTATCCCGCTGATTCTGTGATAACCGTATTTACCGCTTTTGTGAGTGAGCTGA 295
QY 5475 tacgctgcgcgcagcgaacacgcagcgcagcagtcagtcagcgaagcgaagcgaaga 5534
Db 294 TACCGCTCGCCGAGCGGAACGACGCGCAGCGAGTCAGTGAGCGAGGAGCGGAAGA 235
QY 5535 ggcgccaatcagcaacgcctctcccgcggttggcgattcattaatgagcag 5588
Db 234 GCGCCCAATACGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAAATGCGAG 181

RESULT 7
AJ281480/c 1004 bp mRNA linear EST 30-JUN-2000
LOCUS 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION AJ281480
VERSION AJ281480.1 GI:6929360
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
REFERENCE 1 (bases 1 to 1004)
AUTHORS Dimopoulos G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyhofstrasse 1, 69117 Heidelberg, Germany.
LOCATION/Qualifiers
1. .1004
/organism="Anopheles gambiae"
/strain="4A r1"
/db_xref="taxon:7165"

FEATURES
source
```

```
/clone="4A3A-P4G8"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN

Query Match 12.5%; Score 723; DB 9; Length 1004;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 823; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4568 gtgcctcaactgattaagcattgtaactgtcagacccaagtttactcatatatacttaga 4627
Db 825 GTGCCTCACTGATTAGCATTTGCTACTGTCAGACCAAGTTTACTCATATATACTTAGA 766
QY 4628 ttgatttaaaacttcatttttaatttaaaggatctagggtgaagatccttttggataatc 4687
Db 765 TTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATC 706
QY 4688 tcatgacaaaatcccttaacgtgagtttcttcacgtgagcgtcagaccccgtagaaa 4747
Db 705 TCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCGTAAGAA 646
QY 4748 agatcaaaagatcttcttgagatcccttttctgcgcgttaactgctgctgttcaaacaa 4807
Db 645 AGATCAAAAGGATCTTCTTGAGATCCTTTTCTGCGGTAATCTGCTGCTTCAACAA 586
QY 4808 aaaaaacacgcctcaccagcgggtgtgttgcgcgatacaagagctaccacactcttttc 4867
Db 585 AAAAAACCGCTTACCAGCGGTGTTGTTGTCGGATCAAGAGCTACCAACTCTTTTTC 536
QY 4868 cgaagtaactggttcagcagagcagatccaataactgtcctctcagtagccgt 4927
Db 525 CGAAGGTAAGTGGCTTACGAGAGCGCAGATACCAATACTGTTCTCTAGTGTAGCCGT 466
QY 4928 agttaggcaccactcaagaactctgtagcaccgcctacatacctcgtctcgttaatcc 4987
Db 465 AGTTAGGCACGACTTCAAGAACTCTGTAGCACCAGCTACATACCTCGCTCTCTAATCC 406
QY 4988 tgtaccagtgctgctgcagtcgagtcgagtaagtcgtcttaccgggttggactcaagac 5047
Db 405 TGTACCAGTGGCTGCTGCCAGTGGCGGTAAGTCTGCTTACCGGTTGGACTCAAGAC 346
QY 5048 gatagttaccgataagcgacgcgctgggtgaacgggggttcgtgcacacagccca 5107
Db 345 GATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCTGTCANACAGCCCA 286
QY 5108 gcttgagcgaacgactcacccaactgagatcacctacagcgtgagctatgagaagcg 5167
Db 285 GCTTGAGCGAAGGACCTACACCGAAGTACCTACAGCGTGAGCTATGAGAAGCG 236
QY 5168 caacgcttcccgaaggagaaagcgagcaggtatccggttaagcgcaggggtcggaacag 5227
Db 225 CCACGCTTCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGCGGCTGGGAACAG 166
QY 5228 gagagcgcagagggagctccaggggaaacccctggatctctttagctcctgctcggt 5287
Db 165 GAGAGCGCAGGAGGAGCTTCCAGGGGGAACCCCTGGTATCTTTTATAGTCTCTCGGGT 106
QY 5288 ttgcgcacctgacttgagcgtcgatcttttggatgctcgtcagggggcgagcgtat 5347
Db 105 TTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGCGGAGCCTAT 46
QY 5348 ggaataaacgcagcaacgcggcctttttacgggttctcgtgacctttt 5392
```



```
BASE COUNT      258 a      256 c      260 g      265 t
ORIGIN

Query Match      12.0%; Score 691; DB 9; Length 1039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5077 ggctgaacgggggttcgtgcacacagccagcttgagcgaacgacctacacccaactg 5136
DB 1039 GGCTGAACGGGGGTTGCTGCACACAGCCAGCTTGAGCGAAGCAGCTACACCGAAGT 980

QY 5137 agtatactacagcgtgagctatgagaaagcgcacgttcccgaaggagaaaggcgac 5196
DB 979 AGTATCTACAGCGTGAGCTATGAGAAAGCGCCAGCTTCCGAAGGGAGAAAGCGGAC 920

QY 5197 aggtatccggttaagcagcgagggtcgaaacagagagcagcagaggtccaggggga 5256
DB 919 AGGTATCCGTTAAGCGCAGGGTTCGGAACAGAGAGCGCACAGGAGGAGCTTCCAGGGGA 860

QY 5257 aacgcctgggtattttatagtcctgctggggttcgcacacctgtgacttgagcgtcattt 5316
DB 859 AACGCGCTGGTATCTTATAGTCTCTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTT 800

QY 5317 ttgtgatcgtcagggggcgagcctatggaanaagcgcagcagcagcgccctttta 5376
DB 799 TTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAANAACGCCAGCAACGCGGCTTTTAA 740

QY 5377 cggctcctggccttttgctggccttttgcacatgttcttctcgcttatccctgat 5436
DB 739 CGGTTCTTGGCCTTTTGGTGGCCTTTTGTCTACATGTTCTTTCCTGCTTATCCCTGAT 680

QY 5437 tctgtgataaacgctattaccgctttagtgagctgataccctgcgcagccgaagc 5496
DB 679 TCTGTGGATAAACGCTATTACCGCCTTTAGTGAGCTGATACCGCTCGCCGACGCGAAGC 620

QY 5497 accgagcagcagtcagtcagcaggaagcaggaagcgcagcagcagcagcagcagcct 5556
DB 619 ACCGAGCGCAGCAGTCAGTCAGGAGGAAGCGGAAGCGGAGCGCCCAATACGCAACCGCCT 560

QY 5557 ctcccgcgcgttgccgatctaatatgcagctggcagcagcaggtttcccgactgaaa 5616
DB 559 CTCCCCGCGGTTGGCGGATTCATTATGACAGCTGGCAGCAGAGTTTCCCGACTGAAA 500

QY 5617 gcggcagtgagcgaacgaataatgtgagttagctactcattagcagccagcagct 5676
DB 499 CGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAGGCAGCCAGGCT 440

QY 5677 ttacactttatgcttcgcgctgatgtgtgtggaattgtgagcgagataacaatttcac 5736
DB 439 TTACACTTTATGCTTCGGGCTCGTATGTGTGGAATTGTGAGCGGATTAACAATTTCAC 380

QY 5737 acagaaacagctatgacctgattacgcca 5767
DB 379 ACAGAAACAGCTATGACCATGATTACGCCA 349

RESULT 10
LOCUS      B1753192
DEFINITION 603026066F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196255 5',
            mRNA sequence.
ACCESSION  B1753192
VERSION    B1753192.1 GI:15744770
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 780)
REFERENCE  NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL
```

```
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11491 row: d column: 17
            High quality sequence stop: 780.

FEATURES
    source
    Location/Qualifiers
    1..780
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5196255"
        /clone_lib="NIH_MGC_114"
        /lab_host="DH10B"
        /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
        Site_2: EcoRV (destroyed); RNA source anonymous pool of
        male brains, age range 23-27 yo. Library is oligo-dT
        primed and directionally cloned (EcoRV site is destroyed
        upon cloning). Average insert size 1.5 kb. Insert size
        range 1-3 kb. Library is normalized and enriched for
        full-length clones and was constructed by C. Gruber
        (Invitrogen). Research Genetics tracking code 019. Note:
        this is a NIH_MGC Library."
    BASE COUNT      111 a      255 c      260 g      154 t
    ORIGIN

Query Match      11.9%; Score 686; DB 10; Length 780;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 ccccgccagctgtgtgcacagccttcctggctcccccactccttctctctctctc 903
DB 54 CCCCCGAGGTGTGTGTGCACAGCCTTCCCTGGCTTCCCACTCCCTTCCCTCCCTC 113

QY 904 tggctcctgggtgagagagtgagggtgagggcgagagatgcagagctgctggtagc 963
DB 114 TGGCTCTCGGTGAGGAGTGGGGGCTGAGGCGCGGAGGATGCAGAGTGTGGTGAC 173

QY 964 gtgcgtggggcgcgctgctggggcattgcctgaagacccccggggccctgtctctgt 1023
DB 174 GTGCGTGGGGCGCGCTGCGGGGCATTTCGCTGAAGACCCCGGGGCGCTGTCTGTCT 233

QY 1024 ttctgggcatcccccttgcggagccacccatggaccccgcttcctgcccacggag 1083
DB 234 TTCTGGGCATCCCTTTGCGGAGCCACCCTGAGGACCCCGCTGCTTTCGCCACCGAG 293

QY 1084 cccaagcagccttggtcaggggtggtagacgctacaaccttccagagtgctgtctacaa 1143
DB 294 CCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCACAGAGTGTCTGTACCA 353

QY 1144 tatgtgacacccctataccacaggttttgagggcacbgagatgtggaaaccccaacctgag 1203
DB 354 TATGTGACACCCCTATACCCAGGTTTTTGAAGGCACCGAGATGTGAACCCCAACCGGTGAG 413

QY 1204 ctgagcagagactgcctgtacctcaacgtgtgacaccataccccggcgctacatcccc 1263
DB 414 CTGAGCGAGGACTGCTGTGATCTCAACGCTGTGACACCATACATACCCCGGCGCTATATCCCC 473

QY 1264 accctctcctcgtctggatctatgggggtggccttctacagtgggcgctcctcttgagc 1323
DB 474 ACCCTTGTCTCTCTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTTCTCTCTGGAC 533

QY 1324 gtgtacatggcgctctcttctgtgtacagggcgagagactgtgtgtgtccatgaactac 1383
DB 534 GTGTACGATGGCGCTTCTTGTGTACAGGCGGAGAGGACTGTGTGTGTCCATGAAGTAC 593

QY 1384 cgggtgggagccttggcttctcgtggcctgcggggagcagagagcccgggcaatgtg 1443
```

115

Qy	5386	gccttttgctggcccttttgctcacatgctctcttctcctgcttatccctcgattctgtgcat	5445
Db	750	GCCTTTTGCTGGCCCTTTTGCTGCATGTTCTTCTCGCTTATCCCTGATTCGTGGAT	691
Qy	5446	aaccgtattaccgcttttgagtgagctgataccgctgcgcgagccgagcagcagcgcgc	5505
Db	690	AACCGTATTACCGGCTTTGAGTGAGCTGATACCGCTCGCGCAGCCGCAACGACGAGCGC	631
Qy	5506	agcagtgctagtgagcaggaagcagcagcagcagcagcagcagcagcagcagcagc	5565
Db	630	AGCGAGTCAGTGAGCGAGGAAGCGAGAGCGGCCCATACGCAAAACGCCCTCTCCCGCG	571
Qy	5566	cgttgccgagctcattaatgagcgtgagcagcagcagcagcagcagcagcagcagc	5625
Db	570	CGTTGGCGGATTTCATTATATGTCAGCTGGCAGCAGCAGGTTTCCGACTGAGAAAGCGGCGAT	511
Qy	5626	gagcgaacgcaattaatgtagttagctcactcattagcagcagcagcagcagcagcagc	5685
Db	510	GAGCGCAACGCAATTAATGTCAGTGTAGTCTACTCATTAGGCACCCAGGCTTTACACTTT	451
Qy	5686	atgctcccgctcgtatgtgtgtggaattgtgagcagcagcagcagcagcagcagcagc	5745
Db	450	ATGCTCCGGCTCGTATGTTCTGTGGAATTTGTGAGCGGATTAACAAATTCACAGGAAC	391
Qy	5746	agctatgaccatgattacgccca	5767
Db	390	AGCTATGACCATGATTACGCCCA	369

RESULT	12
AL646751/c	
LOCUS	AL646751 XGC-neurula Silurana tropicalis cDNA clone Tneu029j21 5',
DEFINITION	AL646751 680 bp mRNA linear EST 13-DEC-2001
ACCESSION	mRNA sequence.
VERSION	AL646751.1 GI:17654643
KEYWORDS	EST.
SOURCE	western clawed frog.
ORGANISM	Silurana tropicalis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Piploidea;
TITLE	Xenopus tropicalis
JOURNAL	Xenopodinae; Silurana.
COMMENT	1 (bases 1 to 680)
	Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
	Sanger Xenopus tropicalis EST project 2001 (10_2001)
	Unpublished (2001)
	Contact: Huckle E
	Sanger Centre
	Hinxton, Cambridgeshire, CB10 1SA, UK
	Email: trop@sanger.ac.uk
	Sanger Xenopus tropicalis EST project 2001
	TROPICALIS_SEQUENCE_ID: Tneu029j21.sp6
	Sequencing primer: SP6
	This sequence is from a Xenopus Gene Collection (XGC) library
	constructed by Aaron M. Zorn.
FEATURES	Location/Qualifiers
source	1..680
	/organism="Silurana tropicalis"
	/db_xref="taxon:8364"
	/clone="Tneu029j21"
	/clone_lib="XGC-neurula"
	/dev_stage="neurula"
	/lab_host="Escherichia coli DH10B"
	/note="vector: pCS107; Site_1: ECORI; Site_2: NotI; CDNA
	was oligo dt primed from Sug of poly A+ RNA from neurula.
	ECORI-NotI cut CDNA was then ligated into pCS107 with
	ECORI at the 5' end and NotI at the 3' end."
BASE COUNT	144 a 201 c 189 g 146 t
ORIGIN	

Query Match	11.8%	Score 679	DB 9	Length 680
Best Local Similarity	100.0%	Pred. No. 0		

Matches	679;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	4940	acttcaagaactctgtagcaccgctacatacctcgctgtgtaatacctctgttaccagtg	4999						
Db	680	ACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGTAACTCTGTACAGTGG	621						
QY	5000	ctgctgcagtgaggcagataagtgcttcttaccaggggttgagactcaagacagatagttaccgg	5059						
Db	620	CTGCTGCCAGTGGCGATAGTCTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGG	561						
QY	5060	ataagggcagcgctcggggtgaacgggggttcgtgcacacagccagcttggagcgaa	5119						
Db	560	ATAAGCGCGAGCGGTTCGGCTGAACGGGGGTTGTCGACACAGCCACCTTGGAGCGAA	501						
QY	5120	cgacttacaccgaactgagatcacctacagctgagctatgagaagcgccacgcttcccg	5179						
Db	500	CGACCTACACCGAAGTACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG	441						
QY	5180	aagggagaaagcgagcaggtatcgggtaagcggcaagggctcggaacagagagcgcaaga	5239						
Db	440	AAGGGAGAAAGCGGACAGGTATCGGTAAGCGGCGAGGGTCGGAACAGAGAGCGCACGA	381						
QY	5240	ggagagcttcagggggaacacgctggtatctttagtctctcggttccgacctct	5299						
Db	380	GGAGCTTCCAGGGGGAAACGCTGTGTATTTATAGTCTCTGTGCGGTTTCGCCACCTCT	321						
QY	5300	gacttgagcgtcgatttttggatgctcgtcagggggcgagcctatggaaaaacgca	5359						
Db	320	GACTTGAGCGTCGATTTTGTGATGCTCTGACAGGGGCGGAGCCTATGGAAAAAGCGCA	261						
QY	5360	gaaagcggtctttagcgttcctggccttttctgtggttcttgctacatgttcttcc	5419						
Db	260	GCAACGCGGCGCTTTTACGGTTCCTGGGCTTTTGTGCGGCTTTTGTCTACATGTTCTTTC	201						
QY	5420	ctggttaccctgattctgttgataacgctattaccgctttagtgagctgataccg	5479						
Db	200	CTGCGTTATCCCTGATTTCTGTGGATAACCGTATACCGCTTTGAGTGAGTGATACCG	141						
QY	5480	ctgcgcgagcgaacgacgagcagcagctcagtgagcagagagcgaagcgaagcgcc	5539						
Db	140	CTCGCGCAGCGAAGCAGCAGCGAGCGAGTCACTGAGCGAGGAAAGCGAAGAGCGCC	81						
QY	5540	caatcgcaaacgctctcccgcgcttcccgagcttcaattcaatgagctggcagaca	5599						
Db	80	CAATACGCAAAACCGCTCTCCCGCGCGTGTGGCGATTATTAATGACAGTGGCAGCA	21						
QY	5600	ggttcccgactggaaagc	5618						
Db	20	GGTTTCCCGACTGGAAAGC	2						

RESULT 13
LOCUS AL656688/c 707 bp mRNA linear EST 13-DEC-2001
DEFINITION AL656688 XGC-neurula Silurana tropicalis cDNA clone TNeu032h21 5',
mRNA sequence.
ACCESSION AL656688
VERSION AL656688.1 GI:17669128
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Silurana.
Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
AUTHORS
TITLE
COMMENT
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: TNeu032h21.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..707
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu032h21"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCSI07; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCSI07 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 153 a 205 c 195 g 154 t
ORIGIN

Query Match	11.8%;	Score 678;	DB 9;	Length 707;					
Best Local Similarity	100.0%;	Pred. No. 0;							
Matches	678;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	4912	cttctagttagcgttagttagccaccacttcaagaactctgtagcagccgctacatac	4971						
Db	678	CTTCTAGTGTAGCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATAC	619						
QY	4972	ctcgcctctgctaaactcctgttaccagtggtcgcagtgagcagtgagcagtgatcttacc	5031						
Db	618	CTCGCTCTGCTTAATCTGTTTACCACTAGTGCTGCCAGTGGCGGATAGTGTCTTACC	559						
QY	5032	gggttgactcaagacgatatgttaccggataaaggcgcagcggcggcggcggcggcggcgg	5091						
Db	558	GGGTTGGACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTCGGGTGAACGGGGGT	499						
QY	5092	tcgtgcacacgcccagcttgagcgaacgacctacacggaactgagatacctacagcgt	5151						
Db	498	TCGTGCACACAGCCCAAGCTTGGAGCGAAGCAGCTACACCGAACTAGATACCTACAGCGT	439						
QY	5152	ggagctatgaaaaagcgcacgcttcccgaaaggagaaagcgcagcaggtatcccggtaaagc	5211						
Db	438	GAGCTATGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGATATCCGGTTAAGC	379						
QY	5212	ggcagggctcggaacagcagcagcgcgcagcgttccaggggggaaacgcctcgggtatcctt	5271						
Db	378	GGCAGGGTCGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGGGAACGCCCTGGTATCTT	319						
QY	5272	tatagtcctgcgggttctgcacacctctgacttgagcgtcgatcttctgtgagctcgtca	5331						
Db	318	TATAGTCTCTGTCGGGTTTCGCCACCTCTGACTTGAGCTGCGATTTTGTGATGCTCGTCA	259						
QY	5332	ggggggcgagcctatggaaaaacgacgacgacgacgacgacgacgacgacgacgacgacg	5391						
Db	258	GGGGGCGGAGCGCTATGGAAAAACGCCACGACGCGCGCTTTTACGGTCTCTGGCGCTT	199						
QY	5392	tgtcggcctttgtcctcacatgttcttctcctgcgttatccctcgtattctgtggaataaccgt	5451						
Db	198	TGCTGGCCTTTTGCTCACATGTTCTTTCTCGCTTATCCCTGATTCCTGTGGATAACCGT	139						
QY	5452	attacgcctttgagtgagctgataccgctcgcgcagcgcgcgcgcgcgcgcgcgcgcgcg	5511						
Db	138	ATTACGCGCTTTTGTGAGTGTATACCGCTCGCGCAGCGCGAAGCGACCGACGCGAG	79						
QY	5512	tcagtgcgcgaggaagcgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgttgg	5571						
Db	78	TCAGTGAGCGGAGGAGCGGAGAGCGGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGG	19						
QY	5572	ccgattcattaatgcagc	5589						
Db	18	CCGATTTCATTATGACG	1						

```

RESULT 14
AJ281449/c
LOCUS
DEFINITION
4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION
AJ281449
VERSION
AJ281449.1 GI:5929329
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
REFERENCE
AUTHORS
Dimpoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.
and Kafatos, F.C.
TITLE
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE
20300950
COMMENT
Contact: Dimpoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerothstrasse 1, 69117 Heidelberg, Germany.
LOCATION/Qualifiers
1. .800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4D5"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT
203 a 206 c 198 g 193 t
ORIGIN
Query Match 11.8% Score 678; DB 9; Length 800;
Best Local Similarity 99.7% Pred. No. 0;
Matches 778; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4620 acttagattgattataaaacttcatttttaattaaaggatctaggtagaagatccctttt 4679
|||||
Db 800 ACTTTAGATTGATTTAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 741
|||||

Qy 4680 tgataactctatgacccaaatcccttaacgtgagtttcttcactgagcgtcagacc 4739
|||||
Db 740 TGATAATCTCATGACCAAAATCCCTTAACTGAGTTTTCGTTCCACTGAGGTCAGACCC 681
|||||

Qy 4740 cgtgaaagatcaagagattctttagatccctttttctgcgctaactctgctt 4799
|||||
Db 680 CGTAGAAGATCAAGAGATCTTTTGAGATCCCTTTTTCCTGCGCGTAATCTGCTGCTT 621
|||||

Qy 4800 gcaacacacacacacacgcgtaccagcgtggtttgtttgctgcgatacagagctaccac 4859
|||||
Db 620 GCAACAAAAAAACACCGCTACCAAGCGTGGTTTGTTCGGATCAAGAGCTACCAAC 561
|||||

Qy 4860 tcttttccgaaggttaactgcttcacagcagcagcagatacaataactgcttcctagt 4919
|||||
Db 560 TCTTTTCCGAAGGTAACTGCTTCAGCAGAGCGCAGATACCAATACTGTTCTCTAGT 501
|||||

Qy 4920 gtaccgtagttagccaccactcaagaactctgtagcaccgctactacacctgctct 4979
|||||

```

```

Db 500 GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACGCCCTACATACCTCGCTCT 441
Qy 4980 gctaactctgttaccagtggtctgctgcccagtgggcgataaagtcgtctctaccgggttga 5039
|||||
Db 440 GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTACCAGGTTGA 381
|||||
Qy 5040 ctcaagacgatatgtaccggataaagcgcagcggcgcggcgtcgagcggggggttcgtgac 5099
|||||
Db 380 CTCAGACGATAGTTACCGGATAAGCGCGACGGCTTCCAGGGGTAACCGGGGTTCTGTCAC 321
|||||
Qy 5100 acagcccgagcttgagcgaacgacctcacggaactgagatactcacagcgtgagctatg 5159
|||||
Db 320 ACAGCCAGCTTGGAGCGCAACGACCTACCGAACTGAGATACCTACAGCGTGACATAG 261
|||||
Qy 5160 agaagcggccacgcttcccgaaaggagaaagcgcagaggtatccggtaagcgcaggg 5219
|||||
Db 260 AGAAAGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGGT 201
|||||
Qy 5220 cggbaacagagcgcagcagggagcttccaggggggaaacgcctggtatctttatagtc 5279
|||||
Db 200 CGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGTCC 141
|||||
Qy 5280 tctcgggttctcgcaacctctgactgagcgtcgatctgtgctcgtcaggggccc 5339
|||||
Db 140 TGTCCGGTTTCCGACACTCTGACTTGAGCGTCTATTTTGTGATGCTCGTCAAGGGGGCG 81
|||||
Qy 5340 gagcctatggaaaaacgcagcgcgcgcctttttacggcttctcggcttttgcgtggcc 5399
|||||
Db 80 GAGCCTATGGAAACGCGCAGCAACGCGGCTTTTACGGTTCTTGGCTTTTGTGGCC 21
|||||

RESULT 15
AL645114/c
LOCUS
DEFINITION
AL645114 XGC-neurula Silurana tropicalis cDNA clone TNeu017b19 5',
mRNA sequence.
ACCESSION
AL645114
VERSION
AL645114.1 GI:16797239
KEYWORDS
EST.
SOURCE
western clawed frog.
ORGANISM
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
REFERENCE
1 (bases 1 to 724)
Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL
Unpublished (2001)
COMMENT
Contact: Huckle E
Sanger Centre
Hinnton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu017b19.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
LOCATION/Qualifiers
1. .724
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu017b19"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/notes="vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT
162 a 206 c 194 g 162 t
ORIGIN

```

```
Query Match      11.7%; Score 673; DB 9; Length 724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4800 gcaacaaaaaacaccacccgtaccagcggtgttttttgcggatcaagagctaccaac 4859
DB 724 GCANACAAAACACCCGCTACCGCGGTGTTTGTTCGGGATCAAGAGCTACCAAC 665

QY 4860 tttttccgaaggtaactggcttcagcgagcgagataccacaaatactgtcctctagt 4919
DB 664 TCTTTTTCGGAAGTAACCTTTCAGCAGAGCGCAGATACCAAAATACTGTTCTTCTAGT 605

QY 4920 gtacgcgtagttagccacacactcaagaactctgtagcaccgcctacatacctcgtct 4979
DB 604 GTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCTCT 545

QY 4980 gctaactcctgtaccagtgctgctgcagtgccgagataagtcgtcttaccgggttga 5039
DB 544 GCTAATCCTTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTGA 485

QY 5040 ctcaagacgatagttaccggataaagcgcgagcggtcgggctgaacgggggttcgtgcac 5099
DB 484 CTCAAGACGATAGTTACCGGATAAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTGCAC 425

QY 5100 acagcccgacttggagcgaaacacacacacacacacacacacacacacacacacacac 5159
DB 424 ACAGCCCACTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATG 365

QY 5160 agaaagcgccagcgtcccggaaggagaaagcgagcaggtatccggttaagcgcgaggt 5219
DB 364 AGAAGCGCCACGCTTCCGAGAGGGAAGGCGGACAGGTATCCGGTANGCGGCAGGGT 305

QY 5220 cggaaacagagcgacgagggagcttcacagggggaacgcgctgtatctttatagtc 5279
DB 304 CGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTGTATCTTTATAGTCC 245

QY 5280 tctcgggtttccgacactctgacttgagcgtcgatttttgatgctcagggggggcg 5339
DB 244 TGTGGGGTTTCGCCACCTGTACTTGAGCGTTCGATTTTGTGATGCTCGTAGGGGGCG 185

QY 5340 gagcctatggaacacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 5399
DB 184 GAGCCTATGGAACCGCCAGCAACGCGGCCCTTTTACGGTTCCCTGGCCCTTTTGCTGGCC 125

QY 5400 ttttgctcacatgttttctcgtgtatccctgattcctgtggataaaccgtattaccgc 5459
DB 124 TTTTGTCTACATGTTCTTTCCTGCGTTATCCCTGTATCTGTGGATAACCGTATTACCGC 65

QY 5460 ctttgagtgagctgatacagctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5519
DB 64 CTTTGAGTGAGTGATACCGCTCGCGCGCAGCGACGACCGAGCGGCGAGCTCAGTGAG 5

QY 5520 cgag 5523
DB 4 CGAG 1
```

Search completed: August 31, 2002, 19:59:53
Job time: 16402 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:37:24 ; Search time 154.73 Seconds
(without alignments)
9155.099 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgatgcctgcaggtc.....ctatgaccatgattacgccca 5767

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Issued Patents NA.*
- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2637	45.7	8854	3	US-09-053-549-1	Sequence 1, Appli
C 2	2630	45.6	6295	2	US-08-659-206A-4	Sequence 4, Appli
C 3	2628	45.6	4950	3	US-08-789-333F-58	Sequence 58, Appli
C 4	2628	45.6	9639	4	US-09-147-208-26	Sequence 26, Appli
C 5	2533	43.9	4283	1	US-08-343-401A-3	Sequence 3, Appli
C 6	2533	43.9	4283	1	US-08-445-265A-1	Sequence 1, Appli
C 7	2533	43.9	4283	3	US-08-990-442-1	Sequence 1, Appli
C 8	2527	43.8	7566	2	US-08-232-016-23	Sequence 23, Appli
C 9	2527	43.8	7639	2	US-08-232-016-22	Sequence 22, Appli
10	2526	43.8	4713	4	US-09-194-285-7	Sequence 7, Appli
C 11	2526	43.8	4724	4	US-09-194-285-8	Sequence 8, Appli
C 12	2482	43.0	7387	4	US-09-238-356-28	Sequence 28, Appli
C 13	2480	43.0	4045	4	US-08-464-700-54	Sequence 54, Appli
14	2475	42.9	7560	4	US-08-844-274-20	Sequence 20, Appli
C 15	2448	42.4	11284	3	US-08-978-741-5	Sequence 5, Appli
C 16	2418	41.9	8225	4	US-08-793-618-1	Sequence 1, Appli
C 17	2413	41.8	6350	2	US-08-385-335A-8	Sequence 8, Appli
C 18	2398	41.6	3819	4	US-09-042-353-393	Sequence 243, App
C 19	2398	41.6	3819	4	US-08-758-417A-243	Sequence 393, App
C 20	2398	41.6	3881	4	US-09-042-353-369	Sequence 243, App
C 21	2398	41.6	3881	4	US-08-758-417A-217	Sequence 369, App
C 22	2398	41.6	6387	1	US-07-721-775A-1	Sequence 217, App
C 23	2398	41.6	6387	1	US-08-339-658-1	Sequence 1, Appli
24	2367	41.0	19307	3	US-08-836-022A-10	Sequence 10, Appli
25	2367	41.0	19307	4	US-09-427-048A-10	Sequence 10, Appli
26	2357	40.9	4696	2	US-08-929-967-15	Sequence 15, Appli
27	2357	40.9	5158	2	US-08-929-967-16	Sequence 16, Appli

28	2357	40.9	7607	1	US-08-223-616-19	Sequence 19, Appl
29	2357	40.9	7607	5	PCT-US95-04228-19	Sequence 19, Appl
30	2357	40.9	9108	5	PCT-US95-04228-45	Sequence 45, Appl
31	2350	40.7	5639	3	US-09-175-690A-1	Sequence 1, Appl
32	2350	40.7	6306	5	PCT-US94-00658-1	Sequence 1, Appl
33	2347	40.7	3343	6	5453363-2	Patent No. 5453363
34	2299	39.9	4054	3	US-09-098-287A-9	Sequence 9, Appl
35	2248	39.0	4622	4	US-08-509-024-6	Sequence 6, Appl
36	2248	39.0	4622	4	US-09-333-279-6	Sequence 6, Appl
37	2248	39.0	9641	2	US-08-374-483-3	Sequence 3, Appl
38	2242	38.9	11958	4	US-08-927-317-7	Sequence 7, Appl
39	2240	38.8	8299	1	US-08-462-014-2	Sequence 2, Appl
40	2240	38.8	8299	3	US-08-923-137-3	Sequence 3, Appl
41	2240	38.8	8299	4	US-08-973-334-5	Sequence 5, Appl
42	2240	38.8	8299	4	US-09-563-869A-5	Sequence 5, Appl
43	2240	38.8	8509	1	US-08-462-014-1	Sequence 1, Appl
44	2240	38.8	8509	4	US-08-973-334-4	Sequence 4, Appl
45	2240	38.8	8509	4	US-09-563-869A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-053-549-1
; Sequence 1, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nallini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053.549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..3694
; OTHER INFORMATION: /product= "hyFLIB protein"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8854
; OTHER INFORMATION: /note= "Sequence of pCIB5520
; Patent No. 6121521
; OTHER INFORMATION: containing coding sequence for hyFLIB protein"

US-09-053-549-1

Query Match 45.7%; Score 2637; DB 3; Length 8854;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaatgcccgcgtcttttacaacgctcgtagactgggaaacccctgcggttacccaac 3190
|||||
Db 3892 CGAATTCAGTGCCTGCGTCTTTTACAAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAC 3951
QY 3191 ttaatcgccctgacgacacatcccccttcccgagctggcgtaatagcggaagagcccgca 3250
|||||
Db 3952 TTAATCGCCCTGCAGCAGACATCCCCCTTTCCGACAGCTGGCGTAATACGGAAGGCCCGCA 4011
QY 3251 ccgatcgcccttcccacagctgacgagctgaatggcggaatggcgccctgatcgcggtatt 3310
|||||
Db 4012 CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATCGGTATT 4071
QY 3311 ttctccttaacgactctgctgggtatttcaacccgcatatgggtgaactctcagtaacaatct 3370
|||||
Db 4072 TTCCTCCTTAGCATCTGTGCGGTATTTCACACCGCATATGGTGCACCTCTCAGTACAATCT 4131
QY 3371 gctctgatccgcatagttlaagccagcccgacaccccgcaacacccgctgacgcgcccct 3430
|||||
Db 4132 GCTCTGATGCCGATAGTTAAGCCAGCCCGACACCCGCCCAACCCGCTGACGCCCT 4191
QY 3431 gacggcttctgtctcccggcattccgcttacagacaagctgtgacgtctctcgggagct 3490
|||||
Db 4192 GACGGCTTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGAGCT 4251
QY 3491 gcatgtgcagaggttttcacgctcatcccgaaacgcgagacgaaagggccctcgtga 3550
|||||
Db 4252 GCATGTGTCAGAGGTTTTCACCGTCTATCACCGAAGCCGAGACGAAAGGGCCCTCGTGA 4311
QY 3551 taecgctattttataggttaatgctcatataataatggtttctttagacgctcaggtggca 3610
|||||
Db 4312 TAGCCCTATTATTATAGTTAATGTTCATGATAATATAGTTTCTTAGAGCTCAGGTGGCA 4371
QY 3611 cttttcgggaaatgctgcggaacccctattgtttatttttcttaaatatcattcaaaata 3670
|||||
Db 4372 CTTTTCGGGGAAATGTCGCGGAACCCCTATTGTTTATTATTTTCTAAATACATTCAAATA 4431
QY 3671 tctatccgctcatgagacaataacccctgataaactcctaataattgaaagaaaga 3730
|||||
Db 4432 TGTATCCGCTCATGAGACAATACCCCTGATAATGTCTCAATAATATTGAAAGGAAGA 4491
QY 3731 gtatgagtattcaacatttccgtgctgcctctattcccttttttggcgcatitttgccttc 3790
|||||
Db 4492 GTATGAGTATTCAACATTTCCGCTGTCGCCCTTATTCCCTTTTTCGGCATTTTGCCTTC 4551
QY 3791 ctggttttctcaccagaaacgctggtgaaagttaaagatcgtgaagatcagttgggtg 3850
Db 4552 CTGTTTTCCTCACCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCATAGTTGGGTG 4611
QY 3851 cagcagtggtttacatcgaaactggaactcaacagcggtaagatccttgagagtttccg 3910
|||||
Db 4612 CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTGAAGTCCCTTGAGAGTTTTCGCC 4671
QY 3911 ccgaagaacggttttccaaatgtagcaccttttaagttctgtctatggtgcggtattat 3970
Db 4672 CCGAAGACGTTTCCAAATGATGAGACATTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 4731
QY - 3971 ccggtattgacccggggaagagcaactcggctgcgcgcatacactattctcagaatgact 4030
Db 4732 CCGGTATTGACCGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACT 4791
QY 4031 tgggtgagtaactcacacagtcacagaaaacatcttaccgatatggcatgacagtaagagaat 4090
Db 4792 TGGTTGAGTACTCACCCAGTCACAGAAAAGCATCTTACGATGGCATGACAGTAAGAGAT 4851
QY 4091 tatgagtgctgcataaaccatagtgataaacactcggcccaacttactcttcgacaacga 4150
|||||
Db 4852 TATGCACTGCTGCCATTAACCATGAGTGATTAACACTGCGGCCAACTTACTTCTTGACAACGA 4911

QY 4151 tcggaggaaccgaaggagctaaacgcgttttttgcacaacatgggggatcatgttaactcgcc 4210
|||||
Db 4912 TCGGAGGACCGAAGAGGACTAAACGCTTTTTCACAACATGGGGATCATGTAACTCGCC 4971
QY 4211 ttgatcgttgggaaccgagcgtgaatgaagccataccaaacagcagcgtgacaccacga 4270
|||||
Db 4972 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACAAACGACGAGCGTGACACCACGA 5031
QY 4271 tgcctgtagcaattggcaacaacgcttcgcaaaactataaactggcgaaactacttactctag 4330
|||||
Db 5032 TGCCTGTAGCAATGGCAACAACGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAG 5091
QY 4331 cttcccgggcaacaattaaatagactggatggagcgggataaagtgcagaccactctctgc 4390
|||||
Db 5092 CTTCCCGGAACAATAAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC 5151
QY 4391 gctcgcccttcggtcggtgctggtttattgctgataataatcctggagcggtagcggtg 4450
|||||
Db 5152 GCTCGGCCCTTCCGCTGCTGCTGTTTATTGCTGATAAATCTGGAGCGGCTGAGCGTGGT 5211
QY 4451 ctgcggttatctatgacgactgggcccagatggttaagccctcccgtatcgtagttatct 4510
|||||
Db 5212 CTCGCGGTATCATTTGCAGCACCTGGGGCCAGATGCTAAGCCCTCCCGTATCGTAGTTATCT 5271
QY 4511 aacacgccccgggagtcaggcaacttatgataagcgaataagacagatcgcgtgagatgggtg 4570
|||||
Db 5272 ACACGACGGGAGTACAGCAACTATGGATGACAGCAATAGACAGATCGGTGAGATAGGTG 5331
QY 4571 cctcactgattaaagcttgtaactctcagaccgaagtttactcatatactatttagattg 4630
Db 5332 CCFCACTGATTAAAGCAATTTGTAACCTGTACAGCAAGTTTACTCATATATACTATTTAGATTG 5391
QY 4631 atttaaaacttcttttaatttaaaggatctagtggaagatccctttttgataaactca 4690
Db 5392 ATTTAAAACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCCTTTTGTGATAACTCA 5451
QY 4691 tgacaaaatcccttaacgctgaggttcttcactcagcgtcagaccocgtagaaaga 4750
|||||
Db 5452 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAANA 5511
QY 4751 tcaaaagctctcttgagatcccttttctgcgctaatctgctgctctgcaacaaaaa 4810
Db 5512 TCNAAAGATCTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTCTGCAACAAAA 5571
QY 4811 aacacccgctaccagcgtggtttgttgcggatacaagagctaccaactctttttccga 4870
Db 5572 AACCAACCGTACCAGCGGTGTTGTTTGCAGATCAAGAGCTACCAACTCTTTTTCGGA 5631
QY 4871 aggttaactggttcacgagcgcagatataccaaaatactccttcttagttagcgttagt 4930
Db 5632 AGGTAACTGGCTTCAGCAGGCGCAGATACCAATACTGTCTCTCTAGTGTAGCCGTAGT 5691
QY 4931 tagggccacacttcaagaactctgtagcaccgctcatatcctcgtctgcttaactcgt 4990
Db 5692 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCTGCTCTGCTAATCCTGT 5751
QY 4991 taccagtggtcgtcgcagtgccgaatgctgtcttaccgggttggaactcaagcagat 5050
Db 5752 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTACCAGGTTGGACTCAAGACGAT 5811
QY 5051 agttacccgataaaggcgcgctcgggtcgggtgaacggggggttctgtagcacaacagccagct 5110
Db 5812 AGTTACCGGATAAGGCGCAGCGGTGCGGTGAAACGGGGGTTCTGTCACACACAGCCAGCT 5871
QY 5111 tggaggaacgacctacacccaactgagataccatcagcgtgagctatgagaaagccca 5170
Db 5872 TGGAGGAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATGAGAAAAGGCCCA 5931
QY 5171 cgtctcccgaaaggagaaaggcgagcaggtatccggttaagcgcagggctcggaacagggag 5230
Db 5932 CGCTTCCCGAAGGAGAAAGGCGGACAGGTATCCGGTAAGCGCGAGGTTCGGACACAGGAG 5991

Qy	3978	tgacgcggggaagagcaactcgggtcgccgcatcacattctcagaatgacttggtga	4037
Db	5455	tgacgcggcgaagagcaactcgggtcgccgcatacactatttctcagaatgacttggtga	5396
Qy	4038	gtactccagctcacagaaagcatcttacgattggcatgacagtagaagaattatcag	4097
Db	5395	gtactccaccagtCACAGAAGAGCTTTACGGATGGCATGACAGTAGAGAAATTTATCGAG	5336
Qy	4098	tgtcgccataaccaatgagtgaataacactgcgccaaacttactctgacaacgacgcgagg	4157
Db	5335	TGCTGCCATTAACCATTGNGTGATAACACTGCGGCCAACTTACTTCTGACNAGCATCGAGG	5276
Qy	4158	acgaaggagctaacgcctttttgcacaacatagggggatacatgttaactcgccttgatcg	4217
Db	5275	ACCAAGGAGCTAAACGCCTTTTTCACAAACATGGGGATCATGTAACTCGCCTTGATCG	5216
Qy	4218	tgggaaacggagctgaatgaagaccataccaaaacagacagcggtgacaccacgatacctgt	4277
Db	5215	TTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACGAGTGCCCTGT	5156
Qy	4278	agcaatggcaacaaacttgcgcaactataactggcgaaactacttacttagcttcccg	4337
Db	5155	AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGNACTACTTACTTAGCTTCCCG	5096
Qy	4338	gcacaataatagatgagtggggcggaataaaagtgcaggaaactcttgcgctcggc	4397
Db	5095	GCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCCTTCTTCGCGCTCGCG	5036
Qy	4398	ccttcggctcgctggttatctgcatataactgagccggtgaagcgtgggtctcccg	4457
Db	5035	CCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGGG	4976
Qy	4458	tatcttgagcactggggccagatggtgaagccctcccgatcgtagttactacaagac	4517
Db	4975	TATCATTTGCAGCATGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC	4916
Qy	4518	ggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagtgctcact	4577
Db	4915	GGGAGTTCAGGCAACATATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCTCACT	4856
Qy	4578	gattaagcattgtaactgcagacaagtttaactcatatatactttagattgattaaa	4637
Db	4855	GATTAGCATTTGTAAGTGTGTGACACCAAGTTTACTTCATATATATTTAGATTGATTTAA	4796
Qy	4638	actcatactttaaaagatcctagtgaaagatccttttgataactctcatgaccaa	4697
Db	4795	ACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGTATATCTCATGACCA	4736
Qy	4698	aatcccttaacgtgagtttttcgttcactgagcgtaagaccccgtagaaaagatcgaagg	4757
Db	4735	AATCCCTTAACGTGAGTTTTCGTTCCACTTGAGCGTACAGCCCGTAGAAAAGTCAAAGG	4676
Qy	4758	atcttctgagatccctttttctgcggtaatctgcttgcttgcaaacaaaaaacacc	4817
Db	4675	ATCTCTTGAGATCCCTTTTTCGGCGTAATCTGCTGTGCAAAACAAAAAACCCCC	4616
Qy	4818	gctaccagcgttggtttgtttgcgggatacagagctacccaactcttttcggaaggtaac	4877
Db	4615	GCTACCAGCGTGGTTGTTTTCGGGATCAAGAGCTACCAACTCTTTTTCGGAAGGTAA	4556
Qy	4878	tgcttcagcagagcgacataccaaaatactgtccctctagttagccgtagttaggcca	4937
Db	4555	TGGCTTCAGCAGACGCCAGATACCAAAATAGTCTCCTTAGTGTAGCCGTAGTTAGGCCA	4496
Qy	4938	ccacttcaagaactctgtagcacccgctacatacctcgtctgtaataccttgttaccagt	4997
Db	4495	CCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTATATCTGTTACCACT	4436
Qy	4998	ggctgctgccagtggcgataaagtcgttaccgggttgactcaagacgatagtttacc	5057
Db	4435	GGCTGCTGCCAGTGGCGATAGTCGTGCTTACCGGGTTGACTCAAGACGATAGTTACC	4376

RESULT

RESULT 3
US-08-789-333F-58/c

: Sequence 58, Application us/08789333F

; Patent No. 6153380

; GENERAL INFORMATION:

APPLICANT: No. 6153380an, Garry P

APPLICANT: Rothenberg, S. M.

;	TITLE OF INVENTION:	TITLE OF INVENTION:	TITLE OF INVENTION:
;	METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR	METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR	METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
;	EFFECTOR PEPTIDES AND RNA MOLECULES	EFFECTOR PEPTIDES AND RNA MOLECULES	EFFECTOR PEPTIDES AND RNA MOLECULES

; FILE REFERENCE: A642601DJBRMSDSS

; CURRENT APPLICATION NUMBER: US/08/789,333F

; CURRENT FILING DATE: 1997

; PRIOR APPLICATION NUMBER: 08/589,108

;
PRIOR FILLING

; PRIOR APPLICATION NUMBER: 08/589,911

; PRIOR FILLING

; NUMBER OF SEQ ID NOS: 102
: COMPLETE: Date In Nov 2

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59

```

; SEQ ID NO
; LENGTH;

```

; LENGTH: 4950
; TYPE: DNA

```

TYPE: DNA
ORGANISM:

```

; ORGANISM: ALLICICAT sequence
:
: FEATURE:

```

OTHER IN

' NOTATION: DESCRIPTION OF SEQUENCE: RETROVIRAL

OTHER INFORMATION: vector with presentation construct sequence.
US-08-789-333P-58

Query Match 45.6%; Score 2628; DB 3; Length 4950; Best Local Similarity 100.0%; Pred. No. 0; Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy 3137	cactggcgctgttttacacgctgctgactgggaaacccctggcggttaccacaacttaac	3196		
Db 4675	CAC TGGCGCTGCTTTTACAACGCTGCTACTGGGAAACCCCTGGCGTTACCCCAACTTAATC	4616		
Qy 3197	gccttgacacacatcccccctttccagcgtgctgaataatagcgaagacgcgcaccgac	3256		
Db 4615	GCC TTGACGACATCCCCCTTTTCGCCACGCTGGCGTAATAGGAAAGAGCCCGCACCCGATC	4556		
Qy 3257	gcccttcccacacgttgccgacgtgaatggcgactgacgctgctgctgctgctgctg	3316		
Db 4555	GCCCTTCCCAACAGTTGGCGACCTGATGGGAATGGCGCTGATGGGTATTTCTCC	4496		
Qy 3317	ttacgcactctgctgggtatttccacccgcataatggtgcactctcagttacaactcgtctg	3376		
Db 4495	TTACGCATCTGCGGTATTTTACACACCCGATATGCTGCACCTCAGTACAATCTGCTCTG	4436		
Qy 3377	atgcgcatagttaagccagccgcagacacccgcacccgccaacacccgctgacgcgcctgacggg	3436		
Db 4435	ATGCCGATAGTTAAGCCAGCCCGACACCCGCCCAACACCCGCTGACGCGCCCTGACCGG	4376		
Qy 3437	cttgctgctcccgacatccgcttlacagacaaagctgtgacgctccgggagctgcatgt	3496		
Db 4375	CTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCTATG	4316		
Qy 3497	gtcagaggttttcacgcgtcaacacgaacgcgcgagacgaagggccctgctgatacgc	3556		
Db 4315	GTACAGAGTTCACCGTCATCACCGAATCCGCGAGACGAAAGGGCCCTCGTGATACGCC	4256		
Qy 3557	tattttatagtttaagtcatgataataatggtttcttaacgctcagtgagtgacatttc	3616		
Db 4255	TATTTTATAGTAAATGATGATAATATGTTTCTTACAGCTCAGTGCGCACTTTTC	4196		
Qy 3617	ggggaaatgtgcgggaacccctattgtttatttttcaataacatcaaatatgtatc	3676		
Db 4195	GGGGAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGATC	4136		
Qy 3677	gcctatgagacataacccctgataaaatgcttcaataataatgaaaaaggagagta tga	3736		
Db 4135	CGCTCATGAGACAATACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATATGA	4076		
Qy 3737	gtattcaacatttcgctgctgcccctattcccttttttggcgacattttgcccctctgttt	3796		
Db 4075	GTATTCACAATTTCCGTGTGCGCCCTATTCCCTTTTTCGCGCATTTTGCCTTCCTGTTT	4016		
Qy 3797	ttgctacccagaaacgctggtgaaagttaaagatgctggaagatcagttgggtgacagag	3856		
Db 4015	TTGCTACCCAGAACGCTGCTGGAAGTAAAGATGCTGAAGATCAGTTGGTGACACGAG	3956		
Qy 3857	tgggttaactcgaactggtatctcaacagcggtaagatccttgagagtttttcgccccgaag	3916		
Db 3955	TGGGTATACATGGAATGATCTCAACACGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAG	3896		
Qy 3917	aacgttttcccaatgatgacacttttaagtctctgctatgtggcggtattattccgcta	3976		
Db 3895	AACGTTTTCCCAATGATGAGCACTTTTAAAGTTCTCTATGTGGCGGGTATTTATCCCGTA	3836		
Qy 3977	ttgacgcccgggcaagacacccgctgcgcgcatcacactattctcagaaatgactgggtg	4036		
Db 3835	TTGACGCGGGCAAGAGCAACTCGCTCGCCGATACACTATTTCTCAGAAATGACTTGGTTG	3776		
Qy 4037	agctaccagttcagacgaagaacatcttaaggaatggaatgacagtaagagaattatgca	4096		
Db 3775	AGTACTCCACGATCAGAAAAAGCATCTTACGGATGGCATGACATAAGAGAATTAATGCA	3716		
Qy 4097	gttgctgcataaccatgagtataacacactggggccaacttacttctgacaacgactcgag	4156		

Db 3715	GTGCTGCCATAACCATGAGTGATACACTGCGGCCAACTTACTTCTGACAAGATCGGAG	3656		
Qy 4157	gaccgaagagctaaaccgcttttttgcacaacatggtgggatacatgtaactgcgcttgatc	4216		
Db 3655	GACCGAAGAGCTAACCCGCTTTTGTGCACAACATGGGGGATCATGTAACCTGCCTTGATC	3596		
Qy 4217	gttgggaacccggagctgaatgaagccataccaaaacgacgagcgtgacaccacgatacctg	4276		
Db 3595	CTTGGGAACCCGAGCTGAATGAAGCATACCAACGACGAGCGTGACACACGATGCTGTG	3536		
Qy 4277	tagcaatggcaacaacgcttgcacaactattaaactgggcaacttacttacttacttcc	4336		
Db 3535	TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGGGAACCTACTTACTCTAGCTTCC	3476		
Qy 4337	ggcaacaattaatagactggatggaggcggataaaagtgcagagaccactctgcgctcgg	4396		
Db 3475	GGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG	3416		
Qy 4397	ccctccggcctggctgggtttatttgcgtgataaactcggagccgggtgagcgtgggtctcg	4456		
Db 3415	CCCTTCCGGCTGGCTGGTATTATTGCTGATAAATCTGGAGCCGCTGAGGCTGGGTCTCG	3356		
Qy 4457	gtatcattgcagcactggggccagatgtaagccctccgctatcgttagttacttactacaga	4516		
Db 3355	GTATCATTTGCAGCACTGGGGCCAGATGTAAGCCCTCCCGTATCTGATGTTATCTACAGA	3296		
Qy 4517	cgggagtcagggcaactatggaacgaataagacagacatcgtcgtgagatgagctgcctcac	4576		
Db 3295	CGGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGTGCTCAG	3236		
Qy 4577	tgattacgactggtaactgctcagaccagtttactcatatatactttagatctgatttaa	4636		
Db 3235	TGATTAAGCATTTGGTAACCTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTA	3176		
Qy 4637	aacttcatttttaattaaaggatcagtgaaagacatccttttttataatcactacagca	4696		
Db 3175	AACTTCATTTTAAATTAAGGATCTAGTGAAAGATCCTTTTGTATATCTCATGACCA	3116		
Qy 4697	aaatcccttaacgtagtttctccactgagcgtcagacccccctagaanaagatacaag	4756		
Db 3115	AAATCCCTTAAGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATCAAG	3056		
Qy 4757	gatctcttgagatcctttttctgcgcgttaactcgtcgtcgtcgtcgtcgtcgtcgtcgtc	4816		
Db 3055	GATCTTCTTGAGATCCTTTTCTGCGGTAATCTGCTGCTTGCACCAAAAAAACCCAC	2996		
Qy 4817	cgtacacagcgtgtgtgttgcgcgacatcagacacacacacacacacacacacacacac	4876		
Db 2995	CGCTACCAAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA	2936		
Qy 4877	ctggcttcagcagagcgcagataccaaataactgctcctcttagtgcgttagttaggccc	4936		
Db 2935	CTGGCTTCAGCAGAGCGCAGATACCAATACTGCTCTTCTAGTGTAGCGGTAGTAGGCC	2876		
Qy 4937	accacttcaagaactctgtagcaccgcctacatacctcgtcgtcgtcgtcgtcgtcgtcgtc	4996		
Db 2875	ACCACTTCAAGAACCTCTGTAGCACCGCCCTACATACCTCGCTGCTTAATCTCTGTACCA	2816		
Qy 4997	tggctgctgcagtgccgataagtcgtcttaccgggttgagactcaagacagatgctac	5056		
Db 2815	TGGCTGCTGCCAGTGGCCGATAAGTCTGCTTACCGGGTTGGACTCAAGACCATGATTAC	2756		
Qy 5057	cggataagggcagcggctcgggctgaaacggggggttcgtgcacacagcccgacttgagac	5116		
Db 2755	CGGATAAGGCGCAGCGCTCGGCTGACCGGGGGTTCGTGTCACACAGCCCGCTTGGAGC	2696		
Qy 5117	gaacacactacacccgaactgagataccctacacgctgagactatgaaagcgcacgcttc	5176		
Db 2695	GAACACCTTACACCCGAACCTGAGATACCTACAGCGTACGCTATGAGAAAGCGCCAGCTTC	2636		
Qy 5177	ccgaagggagaaagcgcacaggtatcgggtaagcggcgggtcggaaacagagagacaca	5236		

Db 8737 AGTACTCACCAGTCAAGAAAAGCATCTTACGGATGGCATGACGTAAGAGAATTATGCA 8678
QY 4097 gtcgtccataaaccatgagtataacactgcggccaacttactcttgacaaacgatcgagg 4156
Db 8677 GTGCTGCCATAACCATAGATAGTATAACACTGCGGCCAACCTTACTTCTGACAACGATCGGAG 8618
QY 4157 gaccgaagagctaacccgtcttttgcacaacatggggatcatataactcgccttgatc 4216
Db 8617 GACCAAGAGAGCTAACCGCTTTTTTGACAACTGGGGATCATGTAACCTCGCCTTGATC 8558
QY 4217 gttggaaacggagctgaatgaagccataccaaacgaagcgtgacacacagatgcctg 4276
Db 8557 GTTGGGAACCGAGCTGAATGAAGCCATACCAACAGCAGCGGTGACACACGATGCCTG 8498
QY 4277 tagcaatggcaaacagcttgccaaactattaaactgggaaactaacttactctagctccc 4336
Db 8497 TAGCAATGGCAACAACGTTGCCAAACTATTAACTGGCGAACTACTTACTAGCTTCCC 8438
QY 4337 ggcacaatataagactggatggaggcggataaaagtgcagagccacttcgcgctcgg 4396
Db 8437 GGCACAATTAATAGACTGGATGGAGGGGATAAAGTTGCAGAGCACACTTCTGCGCTCGG 8378
QY 4397 cccctccggctggctgtttattgtcgataaaactctggagccggtgagcgtgggtctcg 4456
Db 8377 CCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGTTGAGCGTGGGTCTCGG 8318
QY 4457 gtaacttcagcaactggggccagatggttaagccctccgtagctagttatctacaaga 4516
Db 8317 GTATCATTTGACACACTGGGGCCAGATGTTAAGCCCTCCCGTATCGTAGTTATCTACACA 8258
QY 4517 cgggagtcaggcaactatgatgaacgaataagacagatcgcgtagatagtgccctac 4576
Db 8257 CCGGAGTCAAGCAACTATGGATGAACGAATAAGACAGATCGCTGAGTAGGTGCTCTAC 8198
QY 4577 tgataagcattggttaactgtcagaccagtttactcaataacttttagattgatttaa 4636
Db 8197 TGATAAGCATTGGTAAGTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTAA 8138
QY 4637 aactcatttttaatttaaaagatactaggtagaactccttttgatataatctcatgaaca 4696
Db 8137 AACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCA 8078
QY 4697 aaatcccttaacgtagtttctgtccactgagctcagaccccgtagaaaagatacaag 4756
Db 8077 AATCCCTTAAGCTGAGTTTCTGTTCCACTGAGCGCTCAGACCCCGTAGAAAAGATCAAG 8018
QY 4757 gatctcttgagatcctttttctgcggttaactgtcgtgttgcaacaaaacacccac 4816
Db 8017 GATCTTCTGAGATCCTTTTTTCTGCGGTAATCTGCTGCTTGCAACAAAAAACCCAC 7958
QY 4817 cgtaccagcgggtgtttgtttgcggatcaagagctaccacactcttttcggaaggtaa 4876
Db 7957 CGCTACCAAGCGGTGTTGTTTGGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTA 7898
QY 4877 ctggcttcagcagcgcagataccaaatactgtccttctagtctagcgtagcttaggcc 4936
Db 7897 CTGGGTTTACGACAGCGGAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCC 7838
QY 4937 accacttcaagaactctgtagcacgcgcctacataactcgtcctgtgtaactcctgttacccag 4996
Db 7837 ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAG 7778
QY 4997 tggctgcgcagtgggcgataagtcgtcttaccgggttggaactcaagacgatagttac 5056
Db 7777 TGGCTGCTGCCAGTGGCGATAGTCTGTCTTACCGGGTTGGACTCAAGACCATAGTTAC 7718
QY 5057 cggataagcgcagcgttcgggctaaacgggggttcgtgcacacagcccaacttggagc 5116
Db 7717 CGGATAAGCGCAGCGTGGCGCTGAACGGGGGTTCTGTGCACACAGCCCGCTTGGAGC 7658
QY 5117 gaacgacctacacccgaactgagataccttacagcgtgagcgtatgagaagcgcacgcttc 5176
Db 7657 GAACGACCTACACCGAAGTGAATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC 7598

QY 5177 ccgaaggggaaagcggacaggtatccggttaagcggcagggctcgaacacagagagcga 5236
Db 7597 CGAAAGGGAAAGGCGGACAGGTATCCGGTAAGCGCAGGGTCCGAACAGGAGCGCA 7538
QY 5237 cgaagagcttccaggggaaacgcctgtatcttattatagtcctgtcgggtttcgcacc 5296
Db 7537 CGAGGAGGTTCCAGGGGAAACGCTGTGTATCTTTATAGTCTCTGCGGGTTTCGCCACC 7478
QY 5297 tctgacttgagcgtgatatttttgatgctcgtcagggggggcggagcctcatgaaaaaag 5356
Db 7477 TCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGAANAACG 7418
QY 5357 ccagaaacggccttttttacgggttctcggcttctcggccttttgcctacatgttct 5416
Db 7417 CCAGAAACCGGCCCTTTTACGGTTCTCGCCCTTTTGTGGCCTTTTGTCTACATGTCT 7358
QY 5417 ttccctgcgttatccctgattctgtggataaacgattaccaccccttggatgaactgata 5476
Db 7357 TTCTCGGTTATCCCTGATTCTGTGGATAACGTAATTAACCGCCTTTGAGTGAGCTGATA 7298
QY 5477 ccgctcgcgcagcgcgaacgacgcgagcgcagcgcagtcagtcagtcgaggaagcgaagagc 5536
Db 7297 CGCCTCGCGCAGCGCAACGACGAGCGAGCGAGTCACTGAGCGAGGAAGCGGAAGAGC 7238
QY 5537 gccataacgcgaacgcgcctctcccgcgcttgcggcgttgcggattcatttaagtgcagtcgacg 5596
Db 7237 GCCCAATAGCAAAACCGCCTCTCCCGCGGTTGGCCGATTCAATTAATGAGCTGGCAGC 7178
QY 5597 acagtttccctgcactggaaaagcggcgcagtcagtcagtcgaggaacgaatgaatgaactga 5656
Db 7177 ACAGGTTTCCGACTGGAAAGCGGCGAGTGAGCGCAACGAATTAATGTGAGTTAGCTCA 7118
QY 5657 ctcatagtcacccaggtttacactttatccttcgctcgtatgcttgcgtgtgtgtggaatg 5716
Db 7117 CTCATTAGCACCCAGGCTTTACACTTTTATGCTTCGGCTCGTATGTTGTGTGGAAATG 7058
QY 5717 tgagcggataacaatttcacaggaacacagcgtatgacctatgacctattacc 5764
Db 7057 TGAGCGGATAACAATTTTACAGGAAACAGCTATGACCATGATTACG 7010

RESULT 5

US-08-343-401A-3/c
; Sequence 3, Application US/08343401A
; Patent No. 5661132
; GENERAL INFORMATION:
; APPLICANT: Swain, William F
; APPLICANT: Macklin, Michael D
; APPLICANT: Eriksson, Eloi
; APPLICANT: Andree, Christophe
; TITLE OF INVENTION: Improved Wound Healing
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: PO Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,401A
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 11-229-9103-9


```
Qy 4873 gtaactggcttcagcagagcagacataccaaaatactgtccttcttagtgaagccgtagtta 4932
|
|
|
Db 2525 gtaactggcttcagcagagcagacataccaaaatactgtccttcttagtgaagccgtagtta 2466
|
|
|
Qy 4933 ggcacacacttcaagaactctgtagcagccctacatacctcgtctctgctaatcctgtta 4992
|
|
|
Db 2465 ggcacacacttcaagaactctgtagcagccctacatacctcgtctctgctaatcctgtta 2406
|
|
|
Qy 4993 ccagtgtgtgtccagtggcagataagtcgtgtcttaccgggttggagctcaagacgatag 5052
|
|
|
Db 2405 ccagtgtgtgtccagtggcagataagtcgtgtcttaccgggttggagctcaagacgatag 2346
|
|
|
Qy 5053 ttaccggataaaggcagcagcgtcggtgtgaacgggggttccgtgcacacacagccagcttg 5112
|
|
|
Db 2345 ttaccggataaaggcagcagcgtcggtgtgaacgggggttccgtgcacacacagccagcttg 2286
|
|
|
Qy 5113 gagcagcagactcacaccgaactagaglacctacagcgtgagctatgagaaagcgcacg 5172
|
|
|
Db 2285 gagcagcagactcacaccgaactagaglacctacagcgtgagctatgagaaagcgcacg 2226
|
|
|
Qy 5173 ctccccgaagggaagggcagcaggtatcccgtaagcggcagggctgggaacagggagag 5232
|
|
|
Db 2225 ctccccgaagggaagggcagcaggtatcccgtaagcggcagggctgggaacagggagag 2166
|
|
|
Qy 5233 cgcagagggagcttcagggggaacccctggtatctttagtctcctgctgggttccg 5292
|
|
|
Db 2165 cgcagagggagcttcagggggaacccctggtatctttagtctcctgctgggttccg 2106
|
|
|
Qy 5293 caccctcgaactgagcgtcgattttgtgctcgtcagggggcggagcctatgaaa 5352
|
|
|
Db 2105 caccctcgaactgagcgtcgattttgtgctcgtcagggggcggagcctatgaaa 2046
|
|
|
Qy 5353 aacgcagcaacgcggccttttaccggttctggtcctttgtggtccttttgcctcaatg 5412
|
|
|
Db 2045 aacgcagcaacgcggccttttaccggttctggtccttttgcctcaatg 1986
|
|
|
Qy 5413 ttcttctcgttattccctgattctgtgtgataaacgattaccgctttgagtgagct 5472
|
|
|
Db 1985 ttcttctcgttattccctgattctgtgtgataaacgattaccgctttgagtgagct 1926
|
|
|
Qy 5473 gataccgctcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5532
|
|
|
Db 1925 gataccgctcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1866
|
|
|
Qy 5533 gagcgcccaa tacgcaaacccctctctcccgcgcttggcgtatcattaatgagctgg 5592
|
|
|
Db 1865 gagcgcccaa tacgcaaacccctctctcccgcgcttggcgtatcattaatgagctgg 1806
|
|
|
Qy 5593 cagcagaggttcccgactgaaagcggcgagtagcgcacacgaattaatgagcttag 5652
|
|
|
Db 1805 cagcagaggttcccgactgaaagcggcgagtagcgcacacgaattaatgagcttag 1746
|
|
|
Qy 5653 ctactcattaggcaccacagcgtttacaccttattgcttccggctcgtatgttggtgga 5712
|
|
|
Db 1745 ctactcattaggcaccacagcgtttacaccttattgcttccggctcgtatgttggtgga 1686
|
|
|
Qy 5713 attgtgagcgataacaatttcacacaggaacacagctatgacctatgacgca 5767
|
|
|
Db 1685 attgtgagcgataacaatttcacacaggaacacagctatgacctatgacgca 1631
```

RESULT 6

```
US-08-445-265A-1/c
; Sequence 1, Application US/08445265A
; Patent No. 5697901
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Elof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
```

```
;
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,265A
; FILING DATE:
; CLASSIFICATION: 604
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 110229.91080
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "plasmid DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1250)
; US-08-445-265A-1
```

Query Match 43.9%; Score 2533; DB 1; Length 4283;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3133 aatbaactggcgtgtgtttacaacgtcgtgactgggaaacccctggcgttaccacaatt 3192

|

|

Db 4265 AATTACATGGCGCTGCTTTTACACGCTGCTGACTGGGAAACCCCTGGCGTTACCCAACTT 4206

|

|

Qy 3193 aatcgcttgagacacatcccccttccagctggcgttaatatagaaagagccgcaccc 3252

|

|

Db 4205 AATGCGCTTGAGACACATCCCGCTTTTCGCAGCTGGCGTAATAGGAAAGAGCGCGCAC 4146

|

|

Qy 3253 gatcgcttcccaacagttgcgcagcgtgaatggcggaatggcgctgatgcggtatttt 3312

|

|

Db 4145 GATCGCCCTTCCCAACACTTGGCAGCCCTGAATGGCGAATGGCGCTGATGGGTTATTT 4086

|

|

Qy 3313 ctcttagcgaatctgtgcggtatttcaacacgcgaatatggtgcaactctcagttacaatctgc 3372

|

|

Db 4085 CTCCCTAGCATCTGTGCGGTATTTACACCGCATATGTCACATCTCTCAGTACAAATCTGC 4026

|

|

Qy 3373 tctgtagcgcgatagttaagcagcccgaccccgcaaccccgctgacgcgcctga 3432

|

|

Db 4025 TCTGATGCCGATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA 3966

|

|

Qy 3433 cgggcttctgctcccgccatccgcttacagacaagctgtgacctctccgggagctgc 3492

|

|

Db 3965 CGGGCTTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 3906

|

|

Qy 3493 atgtgtcagaggttttccacgtcatcccgaaacgcgcagacgaagggcctctgtata 3552

|

|

Db 3905 ATGTGTACAGAGGTTTTCACCGCTCATCCGAAACCCGCGAGACGAAAGGGCCCTCGTGATA 3846

|

|

Qy 3553 cgcctattttataggttaagtcatgataataagttcttagacgtcaagtgagcact 3612

|

|

Db 3845 CGCCTATTTTATAGGTTAATGTCATGATAATAAATGGTTTCTTAGACGTCAGGTGGCACT 3786

|

|

Qy 3613 ttccggggaaatgtgcgcggaacccctattgtttattttttaaataacattcaaatg 3672

|

|

Db 3785 TTTCGGGAAATGTGCGCGGAACCCCTATTGTTTATTATTTTCTAANTACATTCANATATG 3726

QY 3673 tatccgctcatagacaaataacccttgataaaatgcttccaataatattgaaaaaggagagt 3732
Db 3725 TATCCGCTCATGAGACAATAAACCTGATAAATGCTTCAATAATATGAAAAAGAGAGT 3666
QY 3733 atgagattacaacatttcgctgcctccttattcccttttttggcgcatttgccttcc 3792
Db 3665 ATGAGATATTTCAACATTTCCGCTGCGCCCTTATCCCTTTTTCGGGCATTTTGCCTTCCT 3606
QY 3793 gtttttgcacccagaaacgctggggaagTaaagatgctgaagatcagttgggtgca 3852
Db 3605 GTTTTTCCTCACCAGAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 3546
QY 3853 cgagtgggttacatogaactggtatccaacagcggtgaagatccttgagagttttccccc 3912
Db 3545 CGAGTGGGTACATCGAATCGAATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC 3486
QY 3913 gaagaacggtttccaatgatgagcaacttttaagttctgctatgagcggtattatcc 3972
Db 3485 GAAGAACGTTTTCCAATGATGACACTTTTAAAGTTCTGCTATGTGGCGGTATTTATCC 3426
QY 3973 cgtattgacgcgggcaagagcaactcggctcgccgatacactattctcagaaatgactg 4032
Db 3425 CGTATTGACCGCGGCAAGAGCAACTCGGTCCGCCGATACATATTCTCAGAAATGACTTG 3366
QY 4033 gttgagtactaccagtcacagaaagcatcttaacggatggcatgacagtaagagaatta 4092
Db 3365 GTTGTAGTACTACCACTCAGAAAAAGCATCTTACGGATGGCATGACATTAAGAGAAATTA 3306
QY 4093 tgcagtgctgcataaccatgagtataacactggtggccaacttaacttctgacaacgac 4152
Db 3305 TGCAGTGTGCCATAACCATGATGATACACTGCGGCCAACCTTACTTTGACAAACGATC 3246
QY 4153 ggaggaccgaagagctaacccctttttgcacaacatggggatcatgtaactccctt 4212
Db 3245 CGAGGACCAAGAGCTAACCCGTTTTTGCACAACTGGGGATCATGTAACTCGCCCTT 3186
QY 4213 gatctgtggaaacccgagctgaatgaagccataaccacaaacgacgagcgtgacaccacgag 4272
Db 3185 GATCGTTGGGAACCGGAGCTGAATGAAGCATACCAACAGGAGCGGTGACACACGATG 3126
QY 4273 cctgtagcaatggcaacacgcttgcccaactattaaactggcgaactacttacttagct 4332
Db 3125 CTGTAGCAATGCCACAACAGTTTGCACAACTATTAACTTGGCGAACTACTTACTTAGCT 3066
QY 4333 tcccgcaacaataatagactggatggaggcggaTaaagttgcaggaccacttctgcgc 4392
Db 3065 TCCGGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACCTTCTGCGC 3006
QY 4393 tcggcccttcggctggctgtttattgtgtataaaactcggagccggtgagcgtgggtct 4452
Db 3005 TCGGCCCTTTCGGCTGGCTGGTTTATTGCTGATAAATCTTGAAGCCGCTGAGCGTGGGTCT 2946
QY 4453 cgcggtatcatcgacactggggccagatgtaagccctcccgatcgtatgtattctac 4512
Db 2945 CCGGCTATCATTCACACTTGGGCGCAGATGGTAAGCCCTCCCGTATCTAGTTATCTAC 2886
QY 4513 acgaggggagtcaggcaactatgatgaacgaaaTlagacagatcgtcagatagtgcc 4572
Db 2885 ACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTCAGATAGGTGCC 2826
QY 4573 tcaactgattaaagcattggttaactgtcaagccaaagtttactatatactatttagattgat 4632
Db 2825 TCACTGATTAAAGCATTTGGTAACCTGTCAGACCAAGTTTACTCATATATATACTTTAGATTGAT 2766
QY 4633 taaaacttcattttaattaaaaggatctaggTgaagatcctttttgataatctcatg 4692
Db 2765 TTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATG 2706
QY 4693 accaaaatcccttaacgtgagtttctgttccactgagcgtcagaccccgtagaaagatc 4752
Db 2705 ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGCTCAGACCCCGTAGAAAAAGATC 2646

QY 4753 aaagatctctttgagatccttttttctgcgcgtaatactgctgcttgcacacacacacacac 4812
Db 2645 AAAGATCTCTTGTAGATCTTTTTTTCGCGTAACTGCTGCTTGCACACACACAAAAA 2586
QY 4813 ccaccctaccagcgtggttggccggaTcaagagactaccacactcttttccgaag 4872
Db 2585 CCACCCTACCAGCGTGGTTGTTTCCCGGATCAAGAGCTACCAACTCTTTTTCGGAAG 2526
QY 4873 gtaactgggttcagcagagcgcagataTaccacaaTactgtccttcttagtgcgtagtta 4932
Db 2525 GTAACGTGGCTTCAGCAGAGCGCAGATACCAATPACTGCTCTTCTAGTAGCCGTAGTTA 2466
QY 4933 ggcacacattcaagaactctgtagcacccgctacatacactcgtcgtctgttaTctgtta 4992
Db 2465 GGCACCACTTCAAGAACTCTGTAGCACGCCCTPACATAGCTCGCTCTGCTAACTCTGTTA 2406
QY 4993 ccagtgcctgcctccagtcgagataagTgcttaccgggttgacacaaacagatag 5052
Db 2405 CCAGTGGCTGCTGCCAGTGGCGATAAGTGTCTTACCGGGTTGGACTCAAGACGATAG 2346
QY 5053 tlaccggataaaggcgcagcggctcggtgaaacgggggttcgtgcacacagccacgcttg 5112
Db 2345 TTACCGGATTAAGCGCAGCGGTGCGGCTGNACGGGGGTTCTGTGCACACAGCCACGTTG 2286
QY 5113 gaggcaacactacacaccgaactgagatacactacagctgtagctaTgaaagcgcacag 5172
Db 2285 GAGCGAACACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCACG 2226
QY 5173 cttccgaaggagaaaggcgaacaggtatccggttaagcggaggtcgaacagagag 5232
Db 2225 CTTCCCGAAGGAGAAAGCGGACAGGTATCCGTTAAGCGGAGGTCGACACGAGAG 2166
QY 5233 cgcagcagggagcttccagggggaaacgcctggtatctttatagtcctgtcggttttcgc 5292
Db 2165 CGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGTCTCTGTCGGGTTTCG 2106
QY 5293 caactcgtactgagcgtcagatttttctgatgctcgtcagggggggggcgcacatgagaa 5352
Db 2105 CACCTCTGACTTGAAGCTGCTGATTTTGTATGCTGCTCAGGGGGGAGGCTTATCGAAA 2046
QY 5353 aacgcgcagcaacgcgccttttacggttccttgcccttttgcgcttttgcacatg 5412
Db 2045 AACGCCAGCAACCGCGCCCTTTTACGGTTCCTGGCCCTTTTCTGGCCCTTTTCTCACATG 1986
QY 5413 ttcttctcgtgttaccctgattctgtggataaacctgattaccgccttttgcgtgagct 5472
Db 1985 TTCTTTCTTCCGTTATCCCTGATTCTGTGGATAACCGTATTACCCTTTTGTAGTAGCT 1926
QY 5473 gatacgcctgcgcgcagcgcgaacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5532
Db 1925 GATACCGCTCGCGCAGCGCAACGACGACGCGCAGCTCAGTGAGCGAGGAAGCGGAA 1866
QY 5533 ggcgcgcacacgcgaacccctctcccgcgcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5592
Db 1865 GAGCGCCCAATACGCCAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGACCTGG 1806
QY 5593 cagcagagtttcccgactggaaagcggcagtcgagcgcgaacgcaatTaatgtgagttag 5652
Db 1805 CACGACAGGTTTCCCGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAG 1746
QY 5653 ctaactcataggcaccccgagcttttaactttatgcttccggcgtcgtatgtgtgga 5712
Db 1745 CTCACATATTAGCACCCCGAGGCTTTTACACTTTATGCTTCCGCTCGTATGTGTGGA 1686
QY 5713 attgtgacgcgataaacaatttcacacagaaacagctatgaccatgattacgca 5767
Db 1685 ATTGTAGCGGATAACAAATTTTACACAGGAACAGCTATGACCATGATTACGCCA 1631

RESULT 7

US-08-990-442-1/c

; Sequence 1, Application US/08990442

; Patent No. 6090790

GENERAL INFORMATION:
: APPLICANT: Eriksson, Elof
: TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESSES:
: ADDRESS: Quarles & Brady
: STREET: 1 South Plinckney Street
: CITY: Madison
: STATE: WI
: COUNTRY: US
: ZIP: 53703
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/990.442
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Berson, Bennett J
: REGISTRATION NUMBER: 37094
: REFERENCE/DOCKET NUMBER: 310558.90028
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 608-251-5000
: TELEFAX: 608-251-9166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4283 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "plasmid DNA"
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(713..721, 981..1250)
: US-08-990-442-1

Query Match 43.9%; Score 2533; DB 3; Length 4283;
Best Local Similarity .99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	3133	aattcaactggccgtgctgttttacaacgctgtagctggggaacccctggcggttacccaactt	3192
Db	4265	AATTCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAACTT	4206
Qy	3193	aatgccttgagacacatcccccttgcagagctggcgtaataagcgaagagcccgacc	3252
Db	4205	AATGCCCTTGACGACATFCCCCCTTTGCCAGCTGGCGTAATAGCGAAGAGCGCGCACC	4146
Qy	3253	gatcgccctcccaacagttgcgagcctgaatggcgaatggcgccctgacggtatttt	3312
Db	4145	GATGCCCTTCCCAACAGTTGCGGAGCCTGAATGCGGAATGGCGCTGATGCGGTATTTT	4086
Qy	3313	ctccttacgcacatctgctggttattcacacgcgcataatggtcaactctcagtacacatcgc	3372
Db	4085	CTCCTTACGCATCTGTGGGATTTTACACCGCATATGGTGCACCTCTCAGTACATCTGC	4026
Qy	3373	tctgatgcgcatagtttaagcagccccgcacacccgcacaccccgctgacgccccctga	3432
Db	4025	TCTGATGCGCATAGTTAAGCGCAGCCCGGACACCGCGCAACACCCGCTGACGCGCCCTGA	3966
Qy	3433	cggactgtctcccgagctccgcttacagacaagctgacgctctcgggagctgc	3492
Db	3965	CGGGCTTCTGTCTCCCGGATCCGCTTACAGCAAGCTGTGACCGTCTCCGGGAGCTGC	3906
Qy	3493	atgtcagaggttttccaccgtcatcacccgaacgcgcgagacgaagggcctcgtgata	3552
Db	3905	ATGTGTACAGGTTTTCACCGTTCATCACCAGAAACGCGCGAGACGAAAGGGCCCTCGTGATA	3846

Qy	3553	cgcctattttataaggttaatgtcatgataataaagggtttcttagacgtcaagtggtgcaact	3612
Db	3845	CGCCTATTTTATAGTTAATGTCATGATAATAAATGGTTCTTAGACGTGAGTGCCACT	3786
Qy	3613	tttcggggaaatgtgcgcgaacccctatttcttatttttctaataacattcaaatatg	3672
Db	3785	TTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTATTTTCTAATACATTCAATATG	3726
Qy	3673	tatccgctcatgagacaataaaccttgataaaatgcttcaataatattgaaaaaggagagt	3732
Db	3725	TATCGCTCATGAGACAATAACCCTGATAAATGCTCAATAATAATTGAAAAAGGAAGAGT	3666
Qy	3733	atgagatttcaacatttcgctgctgccttatttcccttttttgcggcattttgccttctc	3792
Db	3665	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTCGCGCATTTTTCCTTCT	3606
Qy	3793	gttttgcctcaccagaaacgcgtgtaaaagtaaaagatgctgaagatcagttggtgca	3852
Db	3605	GTTTGTGCTACCCAGAAACGCTGTGAAGTAAGATGCTGAAGATCAGTTGGGTGCA	3546
Qy	3853	cgagtgggttacatcgaaactggatctcaacagcggtaagatccttgagagtttttcgccc	3912
Db	3545	CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC	3486
Qy	3913	gaagaacgttttccaatgatgagcacttttaaagtctctgctatgtggcgcggtattatcc	3972
Db	3485	GAAGAACGTTTTCCAAATGATGAGCACATTTTAAAGTTCTGCTATGTGCGCGGTATTATCC	3426
Qy	3973	cgattgacgcgggcaagagcaactcggtgcgcacacacacattcttcagagatgacttg	4032
Db	3425	CGATTGACGCGGGCAAGAGCACTCGTCCGCCCATACACTATTCAGAAATGACTTG	3366
Qy	4033	gttgagtactcaccagtcacagaaagcatcttaccgattggcgatgagcagtaagagaatca	4092
Db	3365	GTTGAGTACTCACAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA	3306
Qy	4093	tgcagtgtgcataaaccatgagtgataaacactgcggcacaacttacttctgcaacagacc	4152
Db	3305	TGCAGTGTGCCATTAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGCAACGATC	3246
Qy	4153	ggagacggaagagagtaaccgctttttgcacaacatgggggcatcatgtaactcgccct	4212
Db	3245	GGAGACCGAAGGAGACTTAACCGCTTTTGTGCACACATGGGGGATCATGTAACTCGCCT	3186
Qy	4213	gactgtgggaacccggagctgaatgaagccataccaaacgcagcagcgtgacacccagatg	4272
Db	3185	GATCGTTGGAAACCGAGCTGAATCAAGCCTAACCAACGACGAGCGTGACACCCAGATG	3126
Qy	4273	cctgtagcaatggcaaacacgttgcgcaaacattataactggcgaaactacttactctagct	4332
Db	3125	CCTGTAGCAATGCCAACACGTTGCGCAAACTTATTAACCTGGCGAACTACTTACTTAGCT	3066
Qy	4333	tcccgccacaacttaataagactggatgagcgcgataaagtgcagggacacactcttcgac	4392
Db	3065	TCCCGGCAACAATTAATAGCTGGATGGAGCGGATAAAGTTGACAGGACACTTCTGCGC	3006
Qy	4393	tcggcccttcgggctggctggtttattgctgataaactcggagccggtgagcgtgggtct	4452
Db	3005	TCGGCCCTTCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGTCT	2946
Qy	4453	cgcggtatcattgcagcactggggccagatggtaagccctcccgctatcgtattatctac	4512
Db	2945	CGCGGTATCATTTGACAGCATGCGGGCCAGATGTTAAAGCCCTCCCGTATCTAGTTACTAC	2886
Qy	4513	acgacgggagtcagcgaactatgataacacgaataagacagatcgtgagatggtgc	4572
Db	2885	ACGACGGGAGTCAGGCAACTATGGATGAACGAATAAGATAGACAGATCGCTGAGATAGTGC	2826
Qy	4573	tcactgattaaacattggttaactgtcagaccagtttactcatatatacttttagattgat	4632
Db	2825	TCACTGATTAAACCATTTGGTAACCTGTGACCAAGTTTACTCATATATATCTTATGATTGAT	2766
Qy	4633	ttaaacacttcatttttaatttaaaaggatcttaggtgaagatccttttttgataatctatg	4692

|||||
Db 2765 TTAATAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATG 2706
QY 4693 accaaatacccttaacgtgagttcttcactgagctcagaccccgtagaagaatc 4752
|||||
Db 2705 ACCAAATCCCTTACGTGAGTTTCGTTCCACTGAGCGTCAGACCCCGTAGAANAAGATC 2646
QY 4753 aaagatctcttgagatcctctttttcttcgctgtaactgctgcttgcacacacacaaa 4812
|||||
Db 2645 AAAGGATCTTCTTGAGATCCTTTTCTGCGGTAATCTGCTGCTTGAACAAAAA 2586
QY 4813 caaccctaccagcggtggtttgtttgacgagatcaagagactaccactcttttccgaag 4872
|||||
Db 2585 CCACCCTACCAAGCGGTGTTTGTTCGCGGATCAAGAGCTACCACTCTTTTTCGGAAG 2526
QY 4873 gtaactggtctcagcagcagcagatcaccaatactgctctctctagttagccgtagtta 4932
|||||
Db 2525 GTAACTGGCTTCAGCAGAGCGCAGATACCAATACATCTCTCTCTAGTAGCGGTAGTTA 2466
QY 4933 ggcacacactcaagaactctgtagcacgcgcctacatacctcgctctgtaactcctgtta 4992
|||||
Db 2465 GGCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTGCTCTGCTTAATCTGTGA 2406
QY 4993 ccagtgctgctgcagtgagcagataagtgctgtcttaccgggttgactcaagaacgtag 5052
|||||
Db 2405 CCAGTGGCTGTGCCAGTGGCGATAAGTCTGTCTTACCGGGTGGACTCAAGACGATAG 2346
QY 5053 tlaccggataaagcgagcagcggtcggtgaaacgggggttcgtgcacacagcccgacttg 5112
|||||
Db 2345 TTACCGGATAAGCGCGACGGTGGCGCTGNACCGGGGGTTCGTGCACACAGCCACGCTTG 2286
QY 5113 gagcgaacgactacacccaactgagatacctacagcgtgagctatgagaagaagcgcaag 5172
|||||
Db 2285 GAGCGAACGACCTTACACCAACTGAGATACCTACAGCGTGAGCTATGAGAAGCGCCACG 2226
QY 5173 cttccgaaggaaggaagcgagcagatccggttaagcggcaggtcggaacaggagag 5232
|||||
Db 2225 CTTCGGAAGGGAAGGCGGACAGGTATCCGGTAAAGCGGCGAGGTGCGAACAGGAGAG 2166
QY 5233 cgacgagggaggtcttcagggggaacgcctgtatctttagtctgctcggtgttcgc 5292
|||||
Db 2165 CGCAGCAGGAGCTTCCAGGGGNAACGCTGGTATCTTTATAGTCTGTCCGGTTTCGC 2106
QY 5293 caactctgactgagcgctgctatcttttgatgctcgtcagggggcgagcctatggaaa 5352
|||||
Db 2105 CACCTCTGACTTGAGCGTCGATTTTGTGTATGCTCTCAGGGGGCGGAGCCTATCGAAA 2046
QY 5353 aacgcagcaacgcgcctttttacggttccctgagccttttgcgttttgccttttgcctacatg 5412
|||||
Db 2045 AAGCCGAGCAACGCGCCCTTTTACGGTTCTTGCCCTTTTGTGGCCTTTTGTCTCATG 1986
QY 5413 ttcttctcgtggttatccctgattctgttgataaacgctattaccgcttttgagtgagct 5472
|||||
Db 1985 TTCTTTTCCGTTAICCCCTGATTCGTGTGGATAACCGTATTACCGCTTTGAGTGAGCT 1926
QY 5473 gatacgcctcgcgcgcgcgaacgacccgagcgcagcgagtcagtgagcgcgagcgcgaa 5532
|||||
Db 1925 GATACCGCTCGCGCAGCGCAAGCAGCGCAGCGCAGCGAGTCACTGAGCGAGGAAAGCGGAA 1866
QY 5533 gagcgcccaatcacgaacgcgcctctcccgcgcgcttgccgaattcattaatgcagctgg 5592
|||||
Db 1865 GAGCGCCCAATACGCAACCGCCTCTCCCGCGCGTGTGCGGATTCATTATGACGCTGG 1806
QY 5593 cagcagaggtttcccgactggaaagcggcgagtgagcgcgaacgaattaatgtgagttag 5652
|||||
Db 1805 CACGACAGGTTTCCCGACTTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGCTAG 1746
QY 5653 ctcactcattagcaccacccagcgttttaactttatgcttccggctcgtatgtgtgtag 5712
|||||
Db 1745 CTCACATATTAGGACCCCGAGGCTTTTACACTTTTATGCTTCCGGCTCGTATGTTGTGGA 1686
QY 5713 attgtgagcggataaacaatttcacagaaacagcgtatgacatgattacgca 5767
|||||

Db 1685 ATTGTGAGCGGATACAAATTTACACAGGAACAGCTATGACCATGATTACGCCA 1631
RESULT 8
US-08-232-016-23/C
; Sequence 23, Application US/08232016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELISSEN, Marc
; APPLICANT: SORTEAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN AARSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; TITLE OF INVENTION: PLANT CELLS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pPS0212
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1785
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated modified bt2 (cryIAb) gene, also designated as t
; OTHER INFORMATION: cryIAb6 gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1793..2026
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA gene 7."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2396..2921
; OTHER INFORMATION: /note= "35S promoter sequence
; OTHER INFORMATION: derived from Cauliflower mosaic virus."
; FEATURE:
; NAME/KEY: misc_feature


```
Db 5312 TGGCTTCAGCAGCGCAGATACCAATACTGCTCTCTAGTGTAGCCGTAGTAGGCCA 5253
QY 4938 ccacttaagaactctgtagcaccgcctacatacctctgcctctgctaatcctgttaccagt 4997
Db 5252 CCATTTCAACAACCTCTGTAGCACCCTGTACATACCTCGCTCTGCTAATCTCTGTACCAGT 5193
QY 4998 ggtgtgtccagtggggataagtcgtgtcttaccggttggaactcaagacagatagttacc 5057
Db 5192 GGCTGTGCCAGTGGCGATAAGTCTGTCTTACCGGTTGGACTCAAGACGATAGTTACC 5133
QY 5058 ggtataggcagcagcgtcgggctgaacggggttctgtgcacacagccagcgttgagcg 5117
Db 5132 GGATAGGCGCAGCGGTGCGGGCTGACGGGGGTTCTGTGCACACAGCCCGCTTGGAGCG 5073
QY 5118 aagcactacacccgaactgagatcacctacagcgtgagctatgagaagcgccacgccttc 5177
Db 5072 AACGACCTACACCGAACTGAGATACCTACAGCTGAGCATTTAGAAAGCGCCACGCTTCC 5013
QY 5178 cgaaggagaaagcgagacaggtatccggttaagcgcaggggtcggaacagagagagcgac 5237
Db 5012 CGAAGGAGAAAGCGGACAGGTATCCGTTAAGCGCAGGGTCCGAACAGGAGAGCGCAC 4953
QY 5238 gagggagctcccgagggaacccctggttatctttagtcctgtcggtttcgccacct 5297
Db 4952 GAGGAGCTTCCAGGGGGAACCCCTGGTATCTTTATAGTCCCTGTGCGGGTTCCGCCACCT 4893
QY 5298 ctgacttgagcgtcgattttgtgctgctgcagggggcgagcctatggaaaaacgc 5357
Db 4892 CTCACCTAGCGTCGATTTTGTGTATCTCTGTCAGGGGGCGGAGCCTATGGAAAAACGC 4833
QY 5358 cagcaacgggcttttaccggttctggttcctgttcttgccttttgcctacatgttctt 5417
Db 4832 CAGCAACGGCGGCTTTTACGGTTCCTGGCCCTTTTGTGCTCACCATGTTCTTT 4773
QY 5418 tctggttatccctgattctgtgataacgtattaccgcttttgagtgagctgatac 5477
Db 4772 TCTGGGTATCCCTGATTCTGTGATAACCGTATTACCGCTTTGAGTGAGCTGATAC 4713
QY 5478 cgtccgcgagcgaacacccagcagcagcagtcagtcagtcagtcagtcagtcagtcagtc 5537
Db 4712 CGCTCCCGCAGCGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4653
QY 5538 cccaatacgaacacccctctcccgcggttgccgattcattatgaatgaatgaatgaatga 5597
Db 4652 CCCAATACGGAACCCGCTCTCCCGCGGCTTGGCCGATTTCATTAATGACGCTGGCACGA 4593
QY 5598 caggttcccgactggaagcggcgagtcgagcgaacgcaattaatgtgagttagctcac 5657
Db 4592 CAGGTTCCCGACTGGNAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC 4533
QY 5658 tcattaggcaccacagctttacactttatgcttcggtcgctgctgctgctgctgctgctg 5717
Db 4532 TCATTAGGCACCCAGGCTTTACACTTTATGCTTCGGCTCGTATGTTGTGGAATTGT 4473
QY 5718 gacggataacaatttcacagaagaacagcgtatgacctgattacgcc 5766
Db 4472 GAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCC 4424
```

RESULT 9

```
US-08-232-016-22/c
; Sequence 22, Application US/08232016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELISSEN, Marc
; APPLICANT: SOETAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN AARSSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; PLANT CELLS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pJD884
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1869
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated bt2 (cryIAb) gene, also designated as the bt884
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1877..2110
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA gene 7."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2480..3005
; OTHER INFORMATION: /note= "35S promoter sequence
; OTHER INFORMATION: derived from Cauliflower mosaic virus."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3006..3665
; OTHER INFORMATION: /note= "Coding sequence of
; OTHER INFORMATION: chloramphenicol acetyl transferase gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3666..4491
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA octopline synthase gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5684..6541
; OTHER INFORMATION: /note= "Sequence complementary to
; OTHER INFORMATION: Patent No. 5952547
; OTHER INFORMATION: the coding sequence of the beta-lactamase gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7155..7639
```

;
OTHER INFORMATION: /note= "TR1' and TR2' promoter
OTHER INFORMATION: derived from Agrobacterium T-DNA."
US-08-232-016-22

Query Match 43.8%; Score 2527; DB 2; Length 7639;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2627; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	3138	actggccgtgctttacaaagctgactgggaaacccctggcgcttaccacaaactaatcg	3197	4098	tgtcgccataaaccatgagtgataaactcgggcaaaacttacttctgacaaagatcgaggg	4157
Db	7136	actggccgtgcttttacaacgtcgtagctgggaaacccctggcgcttaccacaaactaatcg	7077	6176	tgctgccatnaaccatgagtgataaactcgggcaaaacttacttctgacaaagatcgaggg	6117
Qy	3198	ccttgagacacatccccctttccgacgctgggtaataagcagaagagcccgacacgacg	3257	4158	accgaagagctaaaccgcttttttgacacacatgggggatcatgtaactcgcttgatcg	4217
Db	7076	ccttgagacacatcccccttttggcagcgtggcgttaataagcagaagagcccgacacgacg	7017	6116	accgaagagctaaaccgcttttttgacacacatgggggatcatgtaactcgcttgatcg	6057
Qy	3258	cccttcccaacagttgcgacgctgaatggcgaatggcgaatggcgaatggcgaatggcgaat	3317	4218	ttgggaacccggagctgaatgaagccataaccacacagcagcgtgacacacagatccctgt	4277
Db	7016	cccttcccaacagttgcgacgctgaatggcgaatggcgaatggcgaatggcgaatggcgaat	6957	6056	ttgggaacccggagctgaatgaagccataaccacacagcagcgtgacacacagatccctgt	5997
Qy	3318	tacgcactgtgcggtattttcacaccgcataatggtgcactctcagtaacaaactctctga	3377	4278	agcaatggcaacacgcttgccgcaaaactattacatggcgaaactacttacttagctcccg	4337
Db	6956	tacgcactgtgcggtattttcacaccgcataatggtgcactctcagtaacaaactctctga	6897	5996	agcaatggcaacacgcttgccgcaaaactattacatggcgaaactacttacttagctcccg	5937
Qy	3378	tgccgcatagtttaagcagcccccagaccccccgaaccccccgaaccccccgaaccccccga	3437	4338	gcaacaaataatacagctgagtgagcggataaaagtgcagacacacacacacacacacac	4397
Db	6896	tgccgcatagtttaagcagcccccagaccccccgaaccccccgaaccccccgaaccccccga	6837	5936	gcaacaaataatacagctgagtgagcggataaaagtgcagacacacacacacacacacac	4397
Qy	3438	ttgtctgctcccgcatccgctttacagacagctgtgacgctctccggagcgtcatgtg	3497	4398	ccttccggctggctgggtttattgtctgaataaactctggagccggtgagcgtgggtctcggg	4457
Db	6836	ttgtctgctcccgcatccgctttacagacagctgtgacgctctccggagcgtcatgtg	6777	5876	ccttccggctggctgggtttattgtctgaataaactctggagccggtgagcgtgggtctcggg	5817
Qy	3498	tcagagggttttcacgctcaccgcaaacgcgcgagagcaaaaggccctcgtgatacgcct	3557	4458	tatcattgacgacactggggccagatggtaagccctcccgatcgtagtatatacaacgac	4517
Db	6776	tcagagggttttcacgctcaccgcaaacgcgcgagagcaaaaggccctcgtgatacgcct	6717	5816	tatcattgacgacactggggccagatggtaagccctcccgatcgtagtatatacaacgac	5757
Qy	3558	attttttaggttaattgcatgataaataagttttcttagacgctcaggtggcaattttcg	3617	4518	gggagctcagcaactatggatgaacgaataagacagacacacacacacacacacacacac	4577
Db	6716	attttttaggttaattgcatgataaataagttttcttagacgctcaggtggcaattttcg	6657	5756	gggagctcagcaactatggatgaacgaataagacagacacacacacacacacacacacac	5697
Qy	3618	gggaaatgctgcgggaacccctatttggttattttctaaatacattcaaatatgtatcc	3677	4578	gattaagcattggtaactgtcagaccaaagtttactcatatataacttttagattgattaaa	4637
Db	6656	gggaaatgctgcgggaacccctatttggttattttctaaatacattcaaatatgtatcc	6597	5696	gattaagcattggtaactgtcagaccaaagtttactcatatataacttttagattgattaaa	5637
Qy	3678	gctcatgagcaaaataacccctgataaaatgcttcaataatattgaaaggaagagatagag	3737	4638	acttcaatttttaatttaaaagatcagtggaagatcctttttgataatcctatcagacaa	4697
Db	6596	gctcatgagcaaaataacccctgataaaatgcttcaataatattgaaaggaagagatagag	6537	5636	acttcaatttttaatttaaaagatcagtggaagatcctttttgataatcctatcagacaa	5577
Qy	3738	tattcaacatttcgctgctccctattcccttttttgcggcattttgccttccctgtttt	3797	4698	aatcccttaacgctgaggttttcgctccactgagcgcagaccccgtagaaaagatcaagg	4757
Db	6536	tattcaacatttcgctgctccctattcccttttttgcggcattttgccttccctgtttt	6477	5576	aatcccttaacgctgaggttttcgctccactgagcgcagaccccgtagaaaagatcaagg	5517
Qy	3798	tgctcaccagaaacgctggggaagtaaaagatgctgaagatcagttgggtgcacagat	3857	4758	atcttcttgagatcccttttttctgctgctgtaactgctgcttgcacacacacacacacac	4817
Db	6476	tgctcaccagaaacgctggggaagtaaaagatgctgaagatcagttgggtgcacagat	6417	5516	atcttcttgagatcccttttttctgctgctgtaactgctgcttgcacacacacacacac	5457
Qy	3858	gggtttacatcgaaactgatactcaacagcggtaagatccttgagagttttcccccgaaga	3917	4818	gctaccagcgggtggtttgttttgccggatcaagagatcacaactcttttccgaaagtaac	4877
Db	6416	gggtttacatcgaaactgatactcaacagcggtaagatccttgagagttttcccccgaaga	6357	5456	gctaccagcgggtggtttgttttgccggatcaagagatcacaactcttttccgaaagtaac	5397
Qy	3918	acgtttttccaatgatgagcaacttttaaaagtcttgcctatgtggcggttatattcccgat	3977	4878	tggtctcagcagagcgcagataccaaaatctctctctagtcctctagtcgtagcttagtgcca	4937
Db	6356	acgtttttccaatgatgagcaacttttaaaagtcttgcctatgtggcggttatattcccgat	6297	5396	tggtctcagcagagcgcagataccaaaatctctctctagtcctctagtcgtagcttagtgcca	5337
Qy	3978	tgacgccccggaagac	4037	4938	ccacttcaagaactctgtagcaccgcctacatacctcctcctctgcttaactctgtttaccagt	4997
Db	6296	tgacgccccggaagac	6237	5336	ccacttcaagaactctgtagcaccgcctacatacctcctcctcctgcttaactctgtttaccagt	5277
Qy	4038	gtactaccagctcacagaaagacatcttacggatggcatgacagtaagagaaattatcgag	4097	4998	ggctgctccagatggcgataagtcgtcttaccgggttggaactcaagacagatattacc	5057
Db	6236	gtactaccagctcacagaaagacatcttacggatggcatgacagtaagagaaattatcgag	6177	5057	ggctgctccagatggcgataagtcgtcttaccgggttggaactcaagacagatattacc	5217

Db	5096	CGAAGGAGAAAGCGGACAGGTATCCGTATAGCGCAGGTGCGAAACAGAGAGCGCAC	5037
Qy	5238	gagggagcttccagggggaaacgcctggatctttatagtcctctgcgggtttccgcacct	5297
Db	5036	GAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCTGTCGGGTTTCGCCACCT	4977
Qy	5298	ctgacttgagcgtcgatatttttgtatgctcgtcaggggggggagcctcatgaaanaacgc	5357
Db	4976	CTGACTTTAGCGTCGATTTTTTTGTGATGTCGTCGACGGGGCGGAGCCCTATGGA AAAACGC	4917
Qy	5358	cagcaacgcggcctttttacggtctccggccttttgcctggccttttgcctacatgtcttt	5417
Db	4916	CAGCAACGCGGCCCTTTTTTACGGTTCCTGGCCCTTTTTTGTCTGGCCCTTTTGTCTACATGTTCTTT	4857
Qy	5418	tcttggtgtatccccctgattcttgtgataaccgatattaccgccttttgagtgagctgatac	5477
Db	4856	TCCTTGCCTTTATCCCTCGATTCTCTGGATTAACCGTATTAACCGCTTTTGAGTGAGCTGTATAC	4797
Qy	5478	cgtctgcgcgcagcgcgaacacacgcagcgcagtcagtgagcagcaggaagcgaagacgc	5537
Db	4796	CGCTCGCGCGACCGCAACCAACGACCGACGCGACGAGTCACTGAGCGAGGAAGCGGAAGAGCG	4737
Qy	5538	cccaatacgcgaaccgcctctcccgcgcgtttggccgattcattaatcagctggcacga	5597
Db	4736	CCCAATACGAAACCGCTCTCCCGCGGTTGGCGGATTCATTTAATCAGCTGGCACA	4677
Qy	5598	caggtttcccgactggaagcgggcagtgagcgcgaacgcgaattaatgtgattagctcac	5657
Db	4676	CAGGTTTTCCGAGTGGAAACGCGGCGAGTCAGCGCAACGCCAATTAATGTGAGTTAGCTTAC	4617
Qy	5658	tcataggcaccacaggccttcaactttatgcttccgcctcgatgtttgtgtggaattgt	5717
Db	4616	TCATTAGGCACCCAGGCCTTACACTTTATGCTTCGGGCTCGTAGTTGTGTGGAAATGT	4557
Qy	5718	gagcggatacaatttcaacagggaaacagctatgacctattaccgcc	5766
Db	4556	GAGCGGTAACAATTTTACACAGGAACAACAGCTATGACCATGATTTACGCC	4508

```

RESULT 10
US-09-194-285-7
; Sequence 7, Application US/09194285
; Patent No. 6355479
; GENERAL INFORMATION:
; APPLICANT: Webb, Susan R.
; APPLICANT: Wingvist, Ola
; APPLICANT: Karlsson, Lars
; APPLICANT: Peterson, Michael R.
; APPLICANT: Peterson, Per A.
; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSRI 536.1
; CURRENT APPLICATION NUMBER: US/09/194, 285
; CURRENT FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4713
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-194-285-7

```

Query Match 43.8%; Score 2526; DB 4; Length 4713;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps

Db	2086	cactggcgcgtctttacaacgctgtgaactggagaaacctggcgcttaccacaattaatc	2145
Qy	3197	gccttgagcacatcccccttcgccagctggcgtaataagcaagaaggccgcaccgatc	3256
Db	2146	gccttgagcacatcccccttcgccagctggcgtaataagcaagaaggccgcaccgatc	2205
Qy	3257	gcccttcccacagttgcgcagcctgaattggcgaatatggcctgatgcggtattttctcc	3316
Db	2206	gcccttcccacagttgcgcagcctgaattggcgaatatggcctgatgcggtattttctcc	2265
Qy	3317	ttagcatctgtcgggtattttaaacocgcatatgttgtgaactctgaataatctgctctg	3376
Db	2266	ttagcatctgtcgggtattttaaacocgcatatgttgtgaactctgaataatctgctctg	2325
Qy	3377	atgcgcgatagttaaagccagcccgacacccgcccaacacccgctgacgcgccctgacggg	3436
Db	2326	atgcgcgatagttaaagccagcccgacacccgcccaacacccgctgacgcgccctgacggg	2385
Qy	3437	cttgtctgcctcccggaatccgcttacagacaagctgtgacgctctccggagagctgcattg	3496
Db	2386	cttgtctgcctcccggaatccgcttacagacaagctgtgacgctctccggagagctgcattg	2445
Qy	3497	gtcagaggttttacocgtoatcacccgaacgcgcgcagacgaaggccctctgatacgc	3556
Db	2446	gtcagaggttttacocgtoatcacccgaacgcgcgcagacgaaggccctctgatacgc	2505
Qy	3557	tatttttataggttaatgtcatgataataatagttttcttagacgtcagatggcacttttc	3616
Db	2506	tatttttataggttaatgtcatgataataatagttttcttagacgtcagatggcacttttc	2565
Qy	3617	ggggaatgtgcgcgaacccctatttgtttattttctaaaatacatccaatatgtatc	3676
Db	2566	ggggaatgtgcgcgaacccctatttgtttattttctaaaatacatccaatatgtatc	2625
Qy	3677	cgctcatgagacaataaacctcgataaattgcttccaataatttgaaaaagggaagatga	3736
Db	2626	cgctcatgagacaataaacctcgataaattgcttccaataatttgaaaaagggaagatga	2685
Qy	3737	gtatttaacatttcogtgcgccttattcccttttttcggcgcattttgccttccgtttt	3796
Db	2686	gtatttaacatttcogtgcgccttattcccttttttcggcgcattttgccttccgtttt	2745
Qy	3797	tgtctacccagaaaacgcgtgtgaaagtataaagatgcgcgaagatcagttgggtgcacgag	3856
Db	2746	tgtctacccagaaaacgcgtgtgaaagtataaagatgcgcgaagatcagttgggtgcacgag	2805
Qy	3857	tgggttacatcgaaacttgatctccaacagcggtaagatccttgagagttttcgcccccgaag	3916
Db	2806	tgggttacatcgaaacttgatctccaacagcggtaagatccttgagagttttcgcccccgaag	2865
Qy	3917	aacgttttccaatgatgagcacattttaaagttctgctatgtggcgcggtattatcccgta	3976
Db	2866	aacgttttccaatgatgagcacattttaaagttctgctatgtggcgcggtattatcccgta	2925
Qy	3977	ttagcgcggggcaagagcaactcggctgcgcgcatacacattcttcagagaatgacttggtg	4036
Db	2926	ttagcgcggggcaagagcaactcggctgcgcgcatacacattcttcagagaatgacttggtg	2985
Qy	4037	agttactcacagtcacagaaaagcatacttaccgataggcataggagtaagaattaagca	4096
Db	2986	agttactcacagtcacagaaaagcatacttaccgataggcataggagtaagaattaagca	3045
Qy	4097	gtctgcctaataacatgagtataaacacttgcggccaaacttactcttgacaacgatcgag	4156
Db	3046	gtctgcctaataacatgagtataaacacttgcggccaaacttactcttgacaacgatcgag	3105
Qy	4157	gaccgaaggagctaaacgcgttttttgcaacaactgggggatacatgtaactcgcttgatc	4216
Db	3106	gaccgaaggagctaaacgcgttttttgcaacaactgggggatacatgtaactcgcttgatc	3165
Qy	4217	gttgggaacccggagctgaattgaagccataaccaaacgacgagcgtgacaccagatgcctg	4276

```
Db 3166 gttgggaaccgagctgaatgaagccataccaaacgacgagcgtgacaccacgagtcgctg 3225
QY 4277 tagcaatggcaacaacgctgacgaacatttaactggaacactactactactagcttccc 4336
Db 3226 tagcaatggcaacaacgctgacgaacatttaactggaacactactactactagcttccc 3285
QY 4337 ggcacaattaatagactggatgagcgagcgatgaatgctgagagcaccactctgcgctcgg 4396
Db 3286 ggcacaattaatagactggatgagcgagcgatgaatgctgagagcaccactctgcgctcgg 3345
QY 4397 cccctccggctggctgtattatgctgataaaactcggagccggtgagcgtgggtctcgcg 4456
Db 3346 cccctccggctggctgtattatgctgataaaactcggagccggtgagcgtgggtctcgcg 3405
QY 4457 gatacatgacgactggggccagatgtagcagccctccgctatcgtatgattatctacaga 4516
Db 3406 gatacatgacgactggggccagatgtagcagccctccgctatcgtatgattatctacaga 3465
QY 4517 cggggagtcaggccaactatgtagaacaataagacagatcgctgagataggtgctctcac 4576
Db 3466 cggggagtcaggccaactatgtagaacaataagacagatcgctgagataggtgctctcac 3525
QY 4577 tgattaagcattggaactgtcagaccgaatttactcatatatactttagattgatttaa 4636
Db 3526 tgattaagcattggaactgtcagaccgaatttactcatatatactttagattgatttaa 3585
QY 4637 aacttcatttttaattaaaggatctaggtagaagatcccttttggatatactcatgacca 4696
Db 3586 aacttcatttttaattaaaggatctaggtagaagatcccttttggatatactcatgacca 3645
QY 4697 aaatcccttaacgtgagtttcgttccactgagcgtgagaccccgtagaagaatcaag 4756
Db 3646 aaatcccttaacgtgagtttcgttccactgagcgtgagaccccgtagaagaatcaag 3705
QY 4757 gatctcttgagatccctttttctgcgctaaatctgctgtgcgttcgaacaaaaaaaccac 4816
Db 3706 gatctcttgagatccctttttctgcgctaaatctgctgttcgaacaaaaaaaccac 3765
QY 4817 cgtacacagcgttggtttgctgcggaataagagctaccactcttttccgaaaggttaa 4876
Db 3766 cgtacacagcgttggtttgctgcggaataagagctaccactcttttccgaaaggttaa 3825
QY 4877 ctggtcttcagcagacgacagatacaaaactgtctctctagctagtcgtagtagtagcc 4936
Db 3826 ctggtcttcagcagacgacagatacaaaactgtctctctagctagtcgtagtagtagcc 3885
QY 4937 accacttcaagaactctgtagcacccgctacatacctcgtctgctgtaactcgtttaccag 4996
Db 3886 accacttcaagaactctgtagcacccgctacatacctcgtctgctgtaactcgtttaccag 3945
QY 4997 tggctgtcgcagtgccgataagtcgtcttaccgggttgagctacagacgatagtttac 5056
Db 3946 tggctgtcgcagtgccgataagtcgtcttaccgggttgagctacagacgatagtttac 4005
QY 5057 cggataagcgcagcggctcgagctgaacgggggttcgtgcacacagccagcttgagc 5116
Db 4006 cggataagcgcagcggctcgagctgaacgggggttcgtgcacacagccagcttgagc 4065
QY 5117 gaacgacctacacgaactgagatacctacagcgtgagctatgagaaagccacgcttc 5176
Db 4066 gaacgacctacacgaactgagatacctacagcgtgagctatgagaaagccacgcttc 4125
QY 5177 ccgaaggagaaaagcggacaggtatccggttaagcggcaggggtcggaaacagagagcgca 5236
Db 4126 ccgaaggagaaaagcggacaggtatccggttaagcggcaggggtcggaaacagagagcgca 4185
QY 5237 cgagggaagctccagggggaacccctggtatctttatagtcctgctgggttttcgcaacc 5296
Db 4186 cgagggaagctccagggggaacccctggtatctttatagtcctgctgggttttcgcaacc 4245
QY 5297 tctgacttgagcgtgattttttgtagtctcgtcagggggcgagcctatggaaaaaacg 5356
Db 4246 tctgacttgagcgtgattttttgtagtctcgtcagggggcgagcctatggaaaaaacg 4305
```

```
QY 5357 ccagcaacgcgcgccttttaacggttccctggccttttgccttttgcctacatgttct 5416
Db 4306 ccagcaacgcgcgccttttaacggttccctggccttttgccttttgcctacatgttct 4365
QY 5417 ttctgcgcttaccctgattctgtggataaacgctattaccgcttgcgtgagctgata 5476
Db 4366 ttctgcgcttaccctgattctgtggataaacgctattaccgcttgcgtgagctgata 4425
QY 5477 ccgctcgcgcgacgcgaacacgcagcagcagcagcagcagcagcagcagcagcagc 5536
Db 4426 ccgctcgcgcgacgcgaacacgcagcagcagcagcagcagcagcagcagcagcagc 4485
QY 5537 gcccaatacgcgaacacgcgcctctcccgcgcttgcgcgcttgcgcgcttgcgcgacg 5596
Db 4486 gcccaatacgcgaacacgcgcctctcccgcgcttgcgcgcttgcgcgcttgcgcgacg 4545
QY 5597 acaggtttcccgactggaagcggcagcagcagcagcagcagcagcagcagcagcagc 5656
Db 4546 acaggtttcccgactggaagcggcagcagcagcagcagcagcagcagcagcagcagc 4605
QY 5657 ctcataggcaccccgactttacactttatgcttccgctcgtatgttggtggaatcg 5716
Db 4606 ctcataggcaccccgactttacactttatgcttccgctcgtatgttggtggaatcg 5716
QY 5717 tgagcggataaacaatttcacacaggaacacgctatgaccatgattacg 5764
Db 4666 tgagcggataaacaatttcacacaggaacacgctatgaccatgattacg 4713
```

RESULT 11

```
US-09-194-285-8
; Sequence 8, Application US/09194285
; Patent No. 6355479
; GENERAL INFORMATION:
; APPLICANT: Webb, Susan R.
; APPLICANT: Wingqvist, Ola
; APPLICANT: Karlsson, Lars
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSI 536.1
; CURRENT APPLICATION NUMBER: US/09/194,285
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4724
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-194-285-8
```

```
Query Match 43.8%; Score 2526; DB 4; Length 4724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3137 cactggcgcgtgcttttacaacgctgactgggaaaccctggcgttaccacacttaac 3196
Db 2097 cactggcgcgtgcttttacaacgctgactgggaaaccctggcgttaccacacttaac 2156
QY 3197 gccttgagcacatcccccttttcgccaagcgtggcgttaatagcgaagggccgcacgcac 3256
Db 2157 gccttgagcacatcccccttttcgccaagcgtggcgttaatagcgaagggccgcacgcac 2216
QY 3257 gcccttcccaacagtgcgcagcctgaatgcgaatggcgcgtgatttttttcc 3316
Db 2217 gcccttcccaacagtgcgcagcctgaatgcgaatggcgcgtgatttttttcc 2276
```


QY	3317	ttacgcatactgtgcggtattttcacacgcgcataatggtgcactctcagtcacaaatctgctctg	3376
Db	2277	ttacgcatactgtgcggtattttcacacgcgcataatggtgcactctcagtcacaaatctgctctg	2336
QY	3377	atgcgcgcatagtttaagcagccccgcacacccgcacacccgcgcacccgcgtgcagcgccctgcagcgg	3436
Db	2337	atgcgcgcatagtttaagcagccccgcacacccgcacacccgcgcacccgcgtgcagcgccctgcagcgg	2396
QY	3437	cttgtctgtctcccgccatccgctttacagacaaagtctgtgacgcctctccggagagctgcatgt	3496
Db	2397	cttgtctgtctcccgccatccgctttacagacaaagtctgtgacgcctctccggagagctgcatgt	2456
QY	3497	gtcagagggttttcacccgctcatcacccgaacgcgcgcagacgaagagcgcctcgtgatacgc	3556
Db	2457	gtcagagggttttcacccgctcatcacccgaacgcgcgcagacgaagagcgcctcgtgatacgc	2516
QY	3557	tatttttataaggttaatgtcatgataataatgggtttcttagacgctcaggtggcacttttc	3616
Db	2517	tatttttataaggttaatgtcatgataataatgggtttcttagacgctcaggtggcacttttc	2576
QY	3617	ggggaaatgtgcgcggaacccctattgtttatttttcaaatcacattcaaatatatac	3676
Db	2577	ggggaaatgtgcgcggaacccctattgtttatttttcaaatcacattcaaatatatac	2636
QY	3677	cgcctcatgagacaataaacccctgataaatgcttcaataatttgaaaaaggaaagagtatga	3736
Db	2637	cgcctcatgagacaataaacccctgataaatgcttcaataatttgaaaaaggaaagagtatga	2696
QY	3737	gtattcaacatttccgtgtgcgccttattcccttttttgcggcattttgccttccctgttt	3796
Db	2697	gtattcaacatttccgtgtgcgccttattcccttttttgcggcattttgccttccctgttt	2756
QY	3797	ttgcctcacccagaaacgcctgggtgaagataaagatcctgaagatcagttgggtgcacgag	3856
Db	2757	ttgcctcacccagaaacgcctgggtgaagataaagatcctgaagatcagttgggtgcacgag	2816
QY	3857	tgggtttacatcgaaactggatctcaacagcggtaagatccttgagagcttttcgcccgcgaag	3916
Db	2817	tgggtttacatcgaaactggatctcaacagcggtaagatccttgagagcttttcgcccgcgaag	2876
QY	3917	aacggttttccaatgatgagcaacttttaaatgtctgtatgtggcgcggtattatcccgta	3976
Db	2877	aacggttttccaatgatgagcaacttttaaatgtctgtatgtggcgcggtattatcccgta	2936
QY	3977	ttagcgcgcgggcaagagcaactcggctgcgcgcatacacattctcagaatgacttggttg	4036
Db	2937	ttagcgcgcgggcaagagcaactcggctgcgcgcatacacattctcagaatgacttggttg	2996
QY	4037	agtactcacccagctcacagaaaaagcatcttcacggatggcatgacagtaagaataattgca	4096
Db	2997	agtactcacccagctcacagaaaaagcatcttcacggatggcatgacagtaagaataattgca	3056
QY	4097	tgctgtccataaccatgagtgataacacgtgcgcgaacttacttctgacaacagatcggag	4156
Db	3057	tgctgtccataaccatgagtgataacacgtgcgcgaacttacttctgacaacagatcggag	3116
QY	4157	gaccgaaggagctaacccgttttttgcacaaatggggatcatgttaactcgccttgatc	4216
Db	3117	gaccgaaggagctaacccgttttttgcacaaatggggatcatgttaactcgccttgatc	3176
QY	4217	gttgggaaccggagctgaatgaagccataccaaaacgcagcgttgacacaacagatgcctg	4276
Db	3177	gttgggaaccggagctgaatgaagccataccaaaacgcagcgttgacacaacagatgcctg	3236
QY	4277	tagcaatggcacaacacgttgcgcaaaactataactggcgaaacttacttactagcttccc	4336
Db	3237	tagcaatggcacaacacgttgcgcaaaactataactggcgaaacttacttactagcttccc	3296
QY	4337	ggcacaacataatagactgtagaggcggaataaagtgcaggaacacattctgcgctcgg	4396
Db	3297	ggcacaacataatagactgtagaggcggaataaagtgcaggaacacattctgcgctcgg	3356

QY	4397	cccttcggcgctggtggtttattgcgcgataaaatctggagccggtgagcgtgggtctcgcg	4455
DB	3357	cccttcggcgctggtggtttattgcgcgataaaatctggagccggtgagcgtgggtctcgcg	3416
QY	4457	gtaataattgcagcacctggggccagatggtaagccctcccgatcgtagtttatatacaga	4516
DB	3417	gtaataattgcagcacctggggccagatggtaagccctcccgatcgtagtttatatacaga	3476
QY	4517	cggggagtcagcgcaactatggaagaaataagacagatcgcgtgagataggctccac	4576
DB	3477	cggggagtcagcgcaactatggaagaaataagacagatcgcgtgagataggctccac	3536
QY	4577	tgataaagcattggttaactgtcagacaaagtcttactcatatacttttagattgatttaa	4636
DB	3537	tgataaagcattggttaactgtcagacaaagtcttactcatatacttttagattgatttaa	3596
QY	4637	aaacttcatttttaatttaaaagatattagtggaagatcccttttttgataatctcatacca	4696
DB	3597	aaacttcatttttaatttaaaagatattagtggaagatcccttttttgataatctcatacca	3656
QY	4697	aaatcccttaacgtgagtttctgtccactgagcgtcagaccctgagaaaaagataaag	4756
DB	3657	aaatcccttaacgtgagtttctgtccactgagcgtcagaccctgagaaaaagataaag	3716
QY	4757	gactctcttgagatccctttttctgcgcgtaactgtctgttcgaaacaaaaaacccac	4816
DB	3717	gactctcttgagatccctttttctgcgcgtaactgtctgttcgaaacaaaaaacccac	3776
QY	4817	cgtaccagcggctgggtttgtttgcgggataaagagctaccaactcttttccgaaggtaa	4876
DB	3777	cgtaccagcggctgggtttgtttgcgggataaagagctaccaactcttttccgaaggtaa	3836
QY	4877	ctgagcttcagcagcgcagataccaaaactctctctcttagtgtagcgttagtgagcc	4936
DB	3837	ctgggcctcagcagcgcagataccaaaactctctctcttagtgtagcgttagtgagcc	3896
QY	4937	accacttcgaagaactctgtagcaccgcctacaacactcgtctgtcaaatcctctttaccag	4996
DB	3897	accacttcgaagaactctgtagcaccgcctacaacactcgtctgtcaaatcctctttaccag	3956
QY	4997	tggctgctccagtggcgataaagctgttcttaccgggttgagactcaagacgatatgttac	5056
DB	3957	tggctgctccagtggcgataaagctgttcttaccgggttgagactcaagacgatatgttac	4016
QY	5057	cggataagcgcagcggctcgggctgaacggggggttcgtgcacacagccagcttgagc	5116
DB	4017	cggataagcgcagcggctcgggctgaacggggggttcgtgcacacagccagcttgagc	4076
QY	5117	gaacgacctacacggaaactcgagatacctacagcgtgagctatgagaaagccacgcttc	5176
DB	4077	gaacgacctacacggaaactcgagatacctacagcgtgagctatgagaaagccacgcttc	4136
QY	5177	ccgaagggaagaaaggcgcagaggtatccggtaagcgcgcagggctcggaacaggagagcga	5236
DB	4137	ccgaagggaagaaaggcgcagaggtatccggtaagcgcgcagggctcggaacaggagagcga	4196
QY	5237	cagagagcttcacggggaaacgcctgatacttttatagtcctgctgggtttcgcacc	5296
DB	4197	cagagagcttcacggggaaacgcctgatacttttatagtcctgctgggtttcgcacc	4256
QY	5297	tctgacttgagcgtcgatatttttgatgctcgtcgaagggggcggagccctatggaaaaag	5356
DB	4257	tctgacttgagcgtcgatatttttgatgctcgtcgaagggggcggagccctatggaaaaag	4316
QY	5357	ccagaacgcggccttttttacgggttccctggccttttgcctggccttttgcctcatgttct	5416
DB	4317	ccagaacgcggccttttttacgggttccctggccttttgcctggccttttgcctcatgttct	4376
QY	5417	tcttcgcgttatccctgattctgtgataaacgtaattaccgccttttgatgagctgata	5476
DB	4377	tcttcgcgttatccctgattctgtgataaacgtaattaccgccttttgatgagctgata	4436
QY	5477	ccgctcgcgcgcgcgaacgcagccgagcgcagcagtcagtgagcgcgaggaagccgaagagc	5536


```

Db 4437 ccgctcgccgagccggaacccgagcgagcagtgacgtcagtgagcgagaaagc 4496
Qy 5537 gcccaatcacgaaccgcctctcccgcggttgcccgattcataatgcagctggcaacg 5596
Db 4497 gcccaatcacgaaccgcctctcccgcggttgcccgattcataatgcagctggcaacg 4556
Qy 5597 acaggttcccgactgaaagcgagcgagtgagcgcaacgaattaatgtgattagctca 5656
Db 4557 acaggttcccgactgaaagcgagcgagtgagcgcaacgaattaatgtgattagctca 4616
Qy 5657 ctcaatgagcaccacagcctttacactttatgcttcggctcgatgtgtgtggaattg 5716
Db 4617 ctcaatgagcaccacagcctttacactttatgcttcggctcgatgtgtgtggaattg 4676
Qy 5717 tgagcgataacaatttcacacaggaacacgctatgaccatgattacg 5764
Db 4677 tgagcgataacaatttcacacaggaacacgctatgaccatgattacg 4724
```

RESULT 12

```

US-09-238-356-28/c
; Sequence 28, Application US/09238356
; Patent No. 6312683
; GENERAL INFORMATION:
; APPLICANT: Kingsman, et al
; TITLE OF INVENTION: Retroviral Vectors
; FILE REFERENCE: 674523-2006
; CURRENT APPLICATION NUMBER: US/09/238,356
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: PCT/GB/03876
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28
; LENGTH: 7387
; TYPE: DNA
; ORGANISM: Artificial Sequence, primer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(7387)
; OTHER INFORMATION: plasmid vector
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AX003207
; DATABASE ENTRY DATE: 2000-08-24
; RELEVANT RESIDUES: (1)..(7387)
US-09-238-356-28
```

```

Query Match 43.0% Score 2482; DB 4; Length 7387;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2632; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3133 aattcaactggcgctgctttacaacgctgactgggaacacccctggcgttacccaactt 3192
Db 7356 AATTCACCTGGCGCGTCTTTACACGCTCGTGACCTGGGAAACCCCTGGCGTTACCAACTT 7297

Qy 3193 aatcgcttcgagcacatccccctttcgacagctggcgtaataagcgaagcccgcaacc 3252
Db 7296 AATCGCCTTCGAGACACATCCCCCTTTCCGACGCTGGCGTAATAGCGAAGAGCCGCGCAC 7237

Qy 3253 gatcgcccttcccaacagttgcgagcgtgaatggcgaaatggcgccctgatcggtatttt 3312
Db 7236 GATCGCCCTTCCCAACAGTTGCGGAGGCTGANTGGCGAATGGCGCCCTGATCGGTTATTTT 7177

Qy 3313 ctcttaccgatctgtgctggttatttcacacgcgcatagggtgacctctcagtcacatctgc 3372
Db 7176 CTCCTTACGCATCTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAAATCTGC 7117

Qy 3373 tctgatcgccgatagtttaagccagcccccacacccgcacaccccgctgacgcccctga 3432
Db 7116 TCTGATCGGCATAGTTAAGCCAGCCCGCACCCGCAACCCCGCTGACCGCCCTTGA 7057
```

```

Qy 3433 cggccttgctgctcccgccatccgcttacagacaagctgtgacgctctccgggagctgc 3492
Db 7056 CGGCTTGCTGCTGCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGAGTGC 6997

Qy 3493 atgtgtcagaggttttccacgctcatcccgaaacgcgagacgaaagggcctcgtgata 3552
Db 6996 ATGTGTcAGAGGTTTTCACCGTCTATCACCgAAACGGCGAGACGNAAGGGCTCTGTGATA 6937

Qy 3553 cgcctatttttatagggttaatgtcatgataaataatggtttcttagacgtcaggtyggcact 3612
Db 6936 CGCCTATTTTATAGGTTAATGTCTATGATAAATAGTGTTCCTTAGACGTCAGGTGGCACT 6877

Qy 3613 ttccgggaaatgtgctgggaacccctatttgtttattttcttaataataataatg 3672
Db 6876 TTTCCGGGAAATGTGCGGAAACCCCTATTGTGTTATTTTCTAAATACATTCAAATATG 6817

Qy 3673 tatccgctcatgagacaataaaccttgataaatgcttcaataatatgaaaaggaagagt 3732
Db 6816 TATCCGCTCATGAGACAATAACCTTGATAAATGCTTCAATAATTGAAAGGAAGAGT 6757

Qy 3733 atgagttatcaacatttccgtgtgcgacctattcccttttttggcgacctttgacctcct 3792
Db 6756 ATGAGTATTCAACATTTCCGTGTGCGCCTTATTCCCTTTTGGCGCATTTTGCCTTCCT 6697

Qy 3793 gtttttgcctccacgaaacgctgtgtaagtaaaagatgctgaagatcagttgggtgca 3852
Db 6696 GTTTTGTCTACCCAGAAACGCTGTGTAAGATAAAGATGCTGAAGATCAGTTGGGTGCA 6637

Qy 3853 cgagtgggttaactcgaactggatctcaacagcgtaagactccttgagagttttcgcccc 3912
Db 6636 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAGATTCCTTGAGAGTTTTCGCCCC 6577

Qy 3913 gaagaacgctttccaaatgatgagcaacttttaaagttctgctatgtggcgcggtattatcc 3972
Db 6576 GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCC 6517

Qy 3973 cgtattgacgcggcggaagacgactcgtcgcgacatacaactatttcagaatgacttg 4032
Db 6516 CGTATTGACGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAANTGACTTG 6457

Qy 4033 gttgagttactcaccagtcacagaaagcatcttacggatggcatgacagtaagaaatta 4092
Db 6456 GTTGAGTACTCACCAGTCACAGAAAGCATCTACGGATGGCATGCACAGTAAGAGAAATA 6397

Qy 4093 tgcagtgctgccaataccatgagtgataaactcgtggcgaacttacttctetgacaaacgac 4152
Db 6396 TGCAGTGTCTGCATAAACCATGAGTGATAACACTGCGGGCAACTTACTTCTGACAAACGATC 6337

Qy 4153 ggagagcgaagagcctaaccgcttttttgcaacaatgggggcatcatgtaactcgccctt 4212
Db 6336 GGAGAGCGAAGGAGCTAACCGCTTTTGTGCACAACATGGGGGATCATGTAACTCGCCCTT 6277

Qy 4213 gatcgttgggaaacggagctgaatgaagcccatccaaacgacgagcgtgacacacagatg 4272
Db 6276 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACACGATG 6217

Qy 4273 ccttagcaaatggcaacaacgttcgcgaacttaatactggcgaaacttacttacttagct 4332
Db 6216 CCTGTAGCAATGCCAACAACTTGGCAAACTATTAACTTGGCGAACTACTTACTTAGCT 6157

Qy 4333 tcccgcaacaataatgactgagtgagcggaataaagttgcaggaacacttctgcgc 4392
Db 6156 TCCCGCAACAATAATAGACTGGATGGAGCGGGATAAAGTTGCAAGGACCCTTCTTGGCGC 6097

Qy 4393 tcggcccttcggcctggcgtttatttctgataaaatcggagcggtgagcgtgggtct 4452
Db 6096 TCGGCCCTTCGGGCTGGCTGGTTTATGCTGATAAATCTGGAGCGGTGAGCGTGGGTCT 6037

Qy 4453 cgcggttatcttgagcgaactggggccagatggttaagccctcccgctatcgtagttatctac 4512
Db 6036 CGCGGTATCTATGACGACTGGGGCCAGATGTTAAGCCCTCCCGCTATCGTAGTTATCTAC 5977

Qy 4513 acgacggggagtcaggcgaactatggtgaacgaaatagacagatcgctgagatagggtgcc 4572
```

||||| 5976 ACACGGGAGTGCAGCACTATGGATGAACGAATAGACAGATCGCTGAGATAGTGCC 5917
||||| 4573 tcaactgattaaagcattggttaactgtcagaccagtttactcatatatactttagattgat 4632
||||| 5916 TCACTCATTAAGCATTTGGTAACGTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT 5957
||||| 4633 ttaaaacttcatttttaatttaaaagatctagtggaagatcctttttgataatctcatg 4692
||||| 5856 TTAATAACTTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGTATAATCTCATG 5797
||||| 4693 accaaatcccttaacgtagtcttttccactgagctcagaccccgtagaagaatc 4752
||||| 5796 ACCAAATCCCTTAACGTGAGTTTCTGTTCCACTGAGCGTCAGACCCCTAGAAAGATC 5737
||||| 4753 aaaggtactcttgagatcctttttctgcgcgtaactgtcgtcttgcaaacaaaaa 4812
||||| 5736 AAGGATCTCTGAGATCCTTTTCTGCGGTAACTCTGCTGCTGGCAACAAAAA 5677
||||| 4813 ccaccgtaccagcgggtgttgttgcgggatcaagagactcaactcttttccgaag 4872
||||| 5676 CCACCCTACACAGCGTGTGTTGTGCGGATCAAGAGTACCAACTCTTTTCCGAAG 5617
||||| 4873 gtaactggttcagcagcagcagatatacaataactgtctcttagttagccgtagtta 4932
||||| 5616 GTAACGTGGCTTCAGCAGAGCGCAGATACCAATAACTGTCTCTAGTGTAGCGGTAGTTA 5557
||||| 4933 ggcacacttcaagaactctgtagcacgcctacatacctcgtctgctaactcctgtta 4992
||||| 5556 GGCCACCCTCAAGAACCTGTAGCACCGCCTACATACCTCGCTCTGCTATCTGTTA 5497
||||| 4993 caagtgcctgcgcagtgccgataagtcgtgtcttaccgggttggaactcaagacgatag 5052
||||| 5496 CCAGTGGCTGCTGCCAGTGGCATAAGTGTCTTTACGGGTTGGACTCAAGACGATAG 5437
||||| 5053 ttaccgataaagcagcagcgcgtcggcctaagcgggggttcgtgcacacagcccaacttg 5112
||||| 5436 TTACCGGATTAAGCGCGCGGTGCGCTGAACGGGGGTTCGTGACACAGCCACGCTTG 5377
||||| 5113 gagcaacgacttaccacccgaactgagatacctacacgcgtgagctatgagaagcgccacg 5172
||||| 5376 GAGCAACGACCTACACCGAAGTACCTACAGCGTGAGCTATGAGAAAGCGCCACG 5317
||||| 5173 cttccgaaggagaaagcagcaggtatccggtaagcggcagggtcgggaacagagagag 5232
||||| 5316 CTTCGGAAGGGAAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGGTCGGAACAGGAGAG 5257
||||| 5233 cgcacgaggagcttcacggggaaacgcctggtatctttagtctcctcgggtttcgc 5292
||||| 5256 CGCACGAGGAGCTTCAGGGGGAACGGCTGTATCTTTATAGTCTCTCGGGTTTCGC 5197
||||| 5293 cactctgacttgagcgtcgatttttgtgatgctcgtcaggggggcgaggcctatgaaa 5352
||||| 5196 CACCTCTGACTTGAGGTCGATTTTGTGATGCTCTGTCAGGGGGGGAGCCTATCGMA 5137
||||| 5353 acgcgacgaacgcggccttttaaggttccctggccttttgccttttgcctcaatcg 5412
||||| 5136 AACGCCAGAACCGCGCCTTTTACGGTTCTCGGCCCTTTTGTGGCCTTTTGTCTACATG 5077
||||| 5413 ttcttctcgttatccctgattctgtgataacgcgtattaccgccttttaagtgaact 5472
||||| 5076 TTCCTTCTCGGTTATCCCTGATTCGTGTGGATACCGTATTACCGCCTTTTGTAGTGAGCT 5017
||||| 5473 gatacgcctcgcgcagcgcgaacccgagcagcagtcagtcagtcagtcagtcagtcagtcag 5532
||||| 5016 GATACCGCTCGCGCAGCGCAACGACCGAGCGACGAGGTCAGTGAGCGAGGAAGCGAA 4957
||||| 5533 gagcgcccaataacgcaaacgcgcctctcccccgcgcgttgcccgattcaattaaagcagctgg 5592
||||| 4956 GAGCGCCCCAATACGCAACCGCTCTCCCGCGCGCTTCCCGGATTCATTAATGCAGCTGG 4897
||||| 5593 cagcagaggttttcccgactggaaagcggcagtcgagcgaacgaatattagtgagttag 5652
|||||

Db 4896 CACGACAGGTTTCCCGACTCGAANGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAG 4837
QY 5653 ctcaactcattagcaccgccaggctttacacttttatcttccggctcgtatgtgtgtgga 5712
|||||
Db 4836 CTCACTCATTAGCACCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTGTGTGGA 4777
|||||
QY 5713 attgtgagcggatacaaatattccacagagaacagcgtatgaccatgattcacgca 5767
|||||
Db 4776 ATTTGTAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 4722
|||||
RESULT 13
US-08-464-700-54
; Sequence 54, Application US/08464700
; Patent No. 6232458
; GENERAL INFORMATION:
; APPLICANT: WEISS, ANTHONY S
; APPLICANT: MARTIN, STEPHEN L
; TITLE OF INVENTION: SYNTHETIC POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,700
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL6520
; FILING DATE: 22-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL9661
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU93/00655
; FILING DATE: 16-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GHC3USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-464-700-54

Query Match 43.0%; Score 2480; DB 4; Length 4045;
Best Local Similarity 99.9%; Pred No. 0;
Matches 2630; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3135 ttcactggccgtcgttttcaacagtcgtgactggggaacccctggcgttaccaccaacttaa 3194
|||||
Db 1 TTCTAGCGCGCTGTTTACAACGTCGTGACCTGGGAAACCCCTGGCGTTACCCAACTTAA 60
|||||
QY 3195 tgcgcttgagcacatccccctttccagcgtggtggttaatagcgaagagggccgcaccga 3254
|||||

Db 61 TCGCCTTGACGACATCCCCCTTTCCGAGCTGGCGTAATACGGAAGGCCCGCACCGA 120
 Qy 3255 tgcgcccctcccaacagttgcgagcctgaatggcgaaatggcgccctgatgcgtattttct 3314
 Db 121 TCGCCCTTTCCCAACAGTTTCCGACGCTGAATGGCGAATGGCCCTGATGCGGTATTTCT 180
 Qy 3315 ccttacgcatctgtcggtatttccacccgcataatggcgactctcagtaacatgtctc 3374
 Db 181 CTTTACGGCATCTGTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAAATCTGCTC 240
 Qy 3375 tgatccgcagtagttaagccagcccccgcacccgcgcacacccgcctgacgcgcctgacg 3434
 Db 241 TGATGCGCATAGTTAAGCCAGCCCGACACCCCGCCACACCCGCTGACGCGCCCTGACG 300
 Qy 3435 gacttgctgtcccgccatccgcttacagacaagctgtgacccgtctccggagctgcat 3494
 Db 301 GCGTTGTCTGCTCCGGGATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCAT 360
 Qy 3495 gtgtcagaggttttccacgctcatccacccgaaacgcgcgagacgaagggccctgtatagc 3554
 Db 361 GTGTGAGAGGTTTTCCCGCTCATCACCGAAGACCGCGGAGACGAAAGGGCCCTCGTATAG 420
 Qy 3555 cctattttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcacttt 3614
 Db 421 CCTATTTTATAGGTTAATGTCTATCATATAATAGTGTCTTACAGCTCAGGTGGCACTTT 480
 Qy 3615 tccgggaaatgtgcggaacccctatttggttatttttctaaatcacattcaaatatgta 3674
 Db 481 TCGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAATATACATTCAAAATATGTA 540
 Qy 3675 tccgctcatgagacataaccctgataaattgcttcaataaattgaaaggaagagtat 3734
 Db 541 TCCGCTCATGAGACAATTAACCTGATAAATGCTTCAATAATATTGAAAAGGAAGAGTAT 600
 Qy 3735 gagtattcaacattccgtgtgcgcccatttcccttttttggcgcatatttgcctccgt 3794
 Db 601 GAGTATTCAACATTTCCGTGTGCGCCTTATTCCCTTTTGTGGGCATTTTGGCTTCCGT 560
 Qy 3795 ttttctcacccagaaacgctgggtgaaagttaaagatcgtgaagatcagttbggtgcacg 3854
 Db 661 TTTTGTCTCACCCAGAAAGCGTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCACG 720
 Qy 3855 agtgggttacatcgaaactggaactcaacagcggtaagatccttgagagtttttcgcccga 3914
 Db 721 AGTGGTTACATCGAATCGAATCTCAACAGCGGTGAAGTCTTTGAGAGTTTTCGCCCGGA 780
 Qy 3915 agaacgttttccaatgatgagcaccttttaagttctgtctatgtgcgcggtattatcccg 3974
 Db 781 AGAAGCTTTTCCAATGATGAGCACATTTTAAAGTTCTGTATGTGGCGCGGTATTATCCCG 840
 Qy 3975 tattgacccgggcaagagcaactcgggtgcgcgcatacactattctcagaatgacttgg 4034
 Db 841 TATTGACGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGT 900
 Qy 4035 tgagtactcacagtcacagaaaagcatcttacgatatggcagtgagcagaagaattatg 4094
 Db 901 TGAGTACTCACAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATG 960
 Qy 4095 cagtgctgcataaacatgatgataaacactgcgcgcacttactcttgacaacgatcgg 4154
 Db 961 CAGTCTGCCATTAACCATGAGTGATTAACNTGCGGCCCAACTTTACTTCTGACAACCATCGG 1020
 Qy 4155 aggacccgaaggagctaaacgcttttttgcacaacatgggggatacatgtaactcgccttga 4214
 Db 1021 AGGACCGAAGGAGCTTAACCGCTTTTGTGCACAACATGGGGGATCATGTAACCTCGCTTGA 1080
 Qy 4215 tctgtggaaacggagctgaatgaagccataaccacacgacgagcgtgacacacgatgcc 4274
 Db 1081 TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACACGATGCC 1140
 Qy 4275 tgtagcaatggcaaacgctgtgcgaactatttaactgagcgaactactactactagcttc 4334

Db 1141 TGTAGCAATGGCAACAACGTTCCGCCAAACTATTAACTGGCGAACTACTTACTTCTAGCTTC 1200
 Qy 4335 ccggcaacaattaaatagactggatggagcgggataaagttagagaccacttctgcgtc 4394
 Db 1201 CCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACTTCTGCGCTC 1260
 Qy 4395 ggccttccggcgtgctggtttatttgcgtataaactctggagcggtagcgtgggtctcg 4454
 Db 1261 GCGCCTTCGCGCTGGCTGGTTTATTGCTGATAAATCTGGAGCGGTTGAGCGTGGTCTCG 1320
 Qy 4455 cgttatcattgcagcactggggccagatgtgtaagccctcccgtatcgtagtattatctacac 4514
 Db 1321 CGGTATCATTTGCAGCACATGGGGCCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACAC 1380
 Qy 4515 gacggggagtcaggcaactatggaatgaacgaatagacagatcgctgagataggtgcctc 4574
 Db 1381 GACGGGAGTCAGGCAACTATGGATGAAGAAATAGACAGATCGGTAGATAGGTGCCCTC 1440
 Qy 4575 actgattaaacgacttgtaactgtcagacaaagttaactatatactatttagattgattt 4634
 Db 1441 ACTGATTAAAGCATTTGGTAACTGTACAGCAAGTTTACTCATATATACTTTAGATTGATTT 1500
 Qy 4635 aaacttcaatttttaatttaaaggatctaggtgaagatcccttttttgataactctcatgac 4694
 Db 1501 AAACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATGAC 1560
 Qy 4695 caaatcccttaacgtgagtttctgctcaactgagcgtcagaccccgtagaagaatcaa 4754
 Db 1561 CAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAANAAGATCAA 1620
 Qy 4755 aggatcttctgagatccctttttctgctgagtaactgtctgtctgtgcacaaacaaaaacc 4814
 Db 1621 AGGATCTCTTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGTCAAAACAAAAAACCC 1680
 Qy 4815 accgctaccagcgggtgggtttgttgcggatcaagagctaccaactcttttccgaagt 4874
 Db 1681 ACCGCTACCAGCGGTGGTTGTTGCCGATCAAGAGCTACCAACTCTTTTCCGAGGT 1740
 Qy 4875 aactggcttcagcagagcagataccacaaactcttctctagtagtagccgttagtagg 4934
 Db 1741 AACTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGTGTAGCCGTAGTAGG 1800
 Qy 4935 ccacacttcaagaactctgtagaccgcctacatacctcgtctgtctaatcctgttaacc 4994
 Db 1801 CCACCACTTCAAGAATCTGTAGCACCGCTACATACCTCGCTGCTTAATCTGTATACC 1860
 Qy 4995 agtggctgtcgcagtgccgataagtcgtcttaccgggttggactcaagacgatgtt 5054
 Db 1861 AGTGGCTGCTGCCAGTGGCGATTAAGTCTGTCTTACC GGTTGGACTCAAGACGATAGTT 1920
 Qy 5055 accgataagggcagcggcgggtgaaacgggggttcgtgcacacagcccagcttga 5114
 Db 1921 ACCGGATAAGGCGCAGCGTCCGGCTGMAACGGGGGTTCTGTCACACAGCCAGCTTGA 1980
 Qy 5115 ggaacgacctacacggaactgagatacctacagcgtgagctatgagaaagcgaacgct 5174
 Db 1981 GCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCATTTGANAAGGCCACGCT 2040
 Qy 5175 tccgaaaggagaaaggcgacaggtatccggtlaagcggcagggctcggaacagagagcg 5234
 Db 2041 TCCCGAAGGCAANAAGCGGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGGAGCGG 2100
 Qy 5235 cagcaggagcttccaggggggaaacgcgtggtatctttatagctcctgtcgggtttcgcca 5294
 Db 2101 CACGAGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCTGTGCGGTTTCGCCA 2160
 Qy 5295 cctctgacttgagcgtcgatttttctgctcgtcagggggggcgagcgtatgaaaaa 5354
 Db 2161 CCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGGCGGAGGCTTATGGAANA 2220
 Qy 5355 gcccagcaacgcgctttttacggttctcctgctcctttgtgcttcttgcacatgtt 5414
 Db 2221 CCGCAGCAACGCGGCTTTTACGGTTCTTGGCTTTTCTGGCTTTTCTGCATCATGTT 2280


```

|||
Db 8942 TACCGCTGGCGCGCAGCCGACCGAGCGCAGCGAGTCAGTGAGCGAGGAGCGGAAGA 8883
Oy 5535 gcgcccatacgcgaacgcgcctctccccgcgctgttgccgattcattaatgcagctggca 5594
Db 8882 GCGCCCAATAGCGCAAAACCGCCTCTCCCGCGCGTGGCGGATTCATTAAATGCAGCTGGCA 8823
Oy 5595 cgacagggttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagct 5654
Db 8822 CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCT 8763
Oy 5655 cactcattaggcacccccaggctttacactttatgcttccggctcgatgtgtgtggaat 5714
Db 8762 CACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGIGGAAT 8703
Oy 5715 tgtgagcggaatacaatttcacaggaacacagctatgaccatgattacy 5764
Db 8702 TGTGAGCGGATACAAATTTACACAGGAACAGCTATGACCATGATTACG 8653

```

Search completed: August 31, 2002, 22:43:50
 Job time: 19694 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:17:17 ; Search time 824.08 Seconds
(without alignments)
12015.151 Million cell updates/sec

Title: US-09-810-861b-3
Perfect score: 5767
Sequence: 1 agcttcacgtcgcaggtc.....ctatgaccatgattacgcca 5767

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5767	100.0	5767	24	Plasmid pTM034. S
2	2777.6	48.2	5897	20	Plasmid pTGP190-1
3	2637.8	45.7	8854	20	DNA encoding chime
4	2637	45.7	3858	22	Plasmid pRK50 used
5	2637	45.7	4960	22	pCMV-I-Cre-pA vect
6	2637	45.7	7332	22	pCMV-I-beta-pA vec
c 7	2637	45.7	8062	22	Plasmid pRK76 used
c 8	2637	45.7	8153	22	Plasmid pRK74 used
9	2635.4	45.7	5711	22	C31-Int expression

10	2633.2	45.7	9335	20	AAV63734
11	2632.8	45.7	6898	20	AAV63742
12	2631.8	45.6	3753	24	ABA04129
13	2631.8	45.6	3755	24	ABA04130
c 14	2631.8	45.6	4283	18	AAAT86449
c 15	2631.8	45.6	4283	18	AAAT86610
c 16	2631.8	45.6	4283	21	AAAT86610
c 17	2631.8	45.6	4283	21	AAAT86610
c 18	2631.8	45.6	4283	21	AAAT86610
c 19	2631.8	45.6	4283	21	AAAT86610
c 20	2631.8	45.6	4283	21	AAAT86610
c 21	2631.8	45.6	4283	21	AAAT86610
c 22	2631.8	45.6	4283	21	AAAT86610
c 23	2631.8	45.6	4283	21	AAAT86610
c 24	2631.8	45.6	4283	21	AAAT86610
c 25	2631.8	45.6	4283	21	AAAT86610
c 26	2631.8	45.6	4283	21	AAAT86610
c 27	2631.8	45.6	4283	21	AAAT86610
c 28	2631.8	45.6	4283	21	AAAT86610
c 29	2631.8	45.6	4283	21	AAAT86610
c 30	2631.8	45.6	4283	21	AAAT86610
c 31	2631.8	45.6	4283	21	AAAT86610
c 32	2631.8	45.6	4283	21	AAAT86610
c 33	2631.8	45.6	4283	21	AAAT86610
c 34	2631.8	45.6	4283	21	AAAT86610
c 35	2631.8	45.6	4283	21	AAAT86610
c 36	2631.8	45.6	4283	21	AAAT86610
c 37	2631.8	45.6	4283	21	AAAT86610
c 38	2631.8	45.6	4283	21	AAAT86610
c 39	2631.8	45.6	4283	21	AAAT86610
c 40	2631.8	45.6	4283	21	AAAT86610
c 41	2631.8	45.6	4283	21	AAAT86610
c 42	2631.8	45.6	4283	21	AAAT86610
c 43	2631.8	45.6	4283	21	AAAT86610
c 44	2631.8	45.6	4283	21	AAAT86610
c 45	2631.8	45.6	4283	21	AAAT86610

ALIGNMENTS

RESULT 1
AAS17547
ID AAS17547 standard; DNA; 5767 BP.

XX AAS17547;

XX 25-FEB-2002 (first entry)

XX Plasmid pTM034.

XX Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;

KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;

KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;

KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;

XX circular; ds.

XX Synthetic.

XX WO200171014-A2.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08468.

XX 17-MAR-2000; 2000US-190440P.

XX (MORT/) MOR T.

PA (SORE/) SOREQ H.

PA (ARNT/) ARNTZEN C.

PA (MASO/) MASON H.

PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H:

WPI: 2002-055120/07.

Production of a transgenic plant which contains a polynucleotide that encodes a human acetylcholinesterase which upon purification is effective against acetylcholinesterase poisoning -

Claim 11; Page 28-31; 42pp; English.

The invention relates to a method of producing a transgenic plant which contains a polynucleotide that encodes human acetylcholinesterase (AChE) which upon purification is effective against acetylcholinesterase poisoning. The method is used for treating a victim of acetylcholinesterase poisoning by administering a therapeutic amount of a physiologically active human acetylcholinesterase expressed in plant tissue. The extensive use of anticholinesterase pesticides with concurrent accidental poisoning, the threat of chemical warfare and environmental concerns demand the development of effective, inexpensive and stage countermeasures and bioremediation solutions. Prior art methods for treating AChE poisoning have used the muscarinic receptor antagonist atropine and oximes to reactivate the organophosphate(OP)-modified AChE. The reversible carbamate, pyridostigmine bromide has also been used as a prophylactic. However, these conventional treatments have limited effectiveness and serious short and long-term side effects and may result in significant performance deficits and even permanent brain damage. This invention permits the utilisation of cholinesterases to counter-act the toxic effects of anti-cholinergic agents, using transgenic plants for the production of the enzymes is cost effective and the product is stable so that the injected enzymes have the advantage of having a long half-life. The transgenic form of the enzymes are also easy to purify. The present sequence is plasmid pRM034, the pGPVkan derivative construct used in the generation of transgenic tomato plants that constitutively express human AChE.

Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;

Query Match	Score 5767	DB 24	Length 5767
-------------	------------	-------	-------------

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 agcttgcatgcctgcagggtcaacatgggtggagcacgacactctcgtctactccaagaata 60

Db 1 agcttgcatgcctgcaggccaacatgggtggagcacgacactctcgtctactccaagaata 60

Qy 61 tcaagatacagtcacagaagaccagagggctattgagactttcaacaaagggtaatat 120

Db 61 tcaagatacagtcctcagaagaccagagggctattgagacttttcaacaaagggtaat 120

Qy 121 cgggaacctcctcggattccattgccagctatctgtcacttcattcgaaggacagtag 180

Db 121 cgggaaacctcctcggattccattgccagctatctgtcacttcacaaaggacagtag 180

Qy 181 aaaaggaagatggcttctacaaatgccatcattgcgataaaagaaaagcctatcqtccaag 240

Db 181 aaaggaagatgagcttctacaaatgccatcattccgataaaagaaagctatcgttcaag 240

Qy 241 aatgccctctaccgacagtqqtcccaaaatqgacccccacccagagaaacatccgtgaa 300

pb 241 aatgcctctaccgacagtgatcccaaatggaaccccaacacgaaacatcctgaa 300

Qy 301 aaagaaagacattccaaaccacatcttcaaaacaaatgaattgaatgaataaactttcaaca 360

[illegible]

0y 361 aagggttaatatcgggaaacctcctcggaattccatgccagctatctctcacttcactcga 120

361
db 120

[illegible]

421 180

QY 1561 atgcacactgctctcccgccagccggcgctgttccacagggccgtgctgcagagcgg 1620
Db 1561 atgcacactgctctcccgccagccggcgctgttccacagggccgtgctgcagagcgg 1620
QY 1621 gcccccgaatggacctgggacgggtgggca tgggagagggccgtcgagggccacgcag 1680
Db 1621 gcccccgaatggacctgggacgggtgggca tgggagagggccgtcgagggccacgcag 1680
QY 1681 ctggcccacctgtgtgtgtctccaggcggaactgggtgggaa tgcacagagctggta 1740
Db 1681 ctggcccacctgtgtgtgtctccaggcggaactgggtgggaa tgcacagagctggta 1740
QY 1741 gcctcctctggacacgacccagcagactcctgtggaaccagaa tggcagctgtgcct 1800
Db 1741 gcctcctctggacacgacccagcagactcctgtggaaccagaa tggcagctgtgcct 1800
QY 1801 caagaaagcgttctccggttctcctgtgcctgtggtaga tggagacttctcagtgac 1860
Db 1801 caagaaagcgttctcctgtgcctgtggtaga tggagacttctcagtgac 1860
QY 1861 accccagagggccctcatcaacggggagacttccacggcgctgcaggtgctgggtgtg 1920
Db 1861 accccagagggccctcatcaacggggagacttccacggcgctgcaggtgctgggtgtg 1920
QY 1921 gtgaaggaatgagggctctatttctgttttacggggcccgagcttcagaaagacaac 1980
Db 1921 gtgaaggaatgagggctctatttctgttttacggggcccgagcttcagaaagacaac 1980
QY 1981 gagtctctcatagccggcgagttcctggcggggtgcgggtgcgggttccccaggta 2040
Db 1981 gagtctctcatagccggcgagttcctggcggggtgcgggtgcgggttccccaggta 2040
QY 2041 agtgccttggaacggcgaggtgtgtgtcctgaattacacagactggtgcattccccagagac 2100
Db 2041 agtgccttggaacggcgaggtgtgtgtcctgaattacacagactggtgcattccccagagac 2100
QY 2101 ccggcacgccctgagggagggccctgagcagatgtgtgtggcgacaccaaatgtcgtgtcccc 2160
Db 2101 ccggcacgccctgagggagggccctgagcagatgtgtgtggcgacaccaaatgtcgtgtcccc 2160
QY 2161 tggggccagctgggtggcgactgggtcccaagggtgcccgggtctacgacctacgtcttt 2220
Db 2161 tggggccagctgggtggcgactgggtcccaagggtgcccgggtctacgacctacgtcttt 2220
QY 2221 gaacacgggttccacgctctctggcccctgtggatgggggtgcccacagggctacagag 2280
Db 2221 gaacacgggttccacgctctctggcccctgtggatgggggtgcccacagggctacagag 2280
QY 2281 atcgagttcatctttgggatccccctgacccctctcgaaactctcgaaactacacggcagagagaaa 2340
Db 2281 atcgagttcatctttgggatccccctgacccctctcgaaactctcgaaactacacggcagagagaaa 2340
QY 2341 atcttcgcccagcgactgatgcga tacttgggccaactttgcccgcacaggggatcccaat 2400
Db 2341 atcttcgcccagcgactgatgcga tacttgggccaactttgcccgcacaggggatcccaat 2400
QY 2401 gagccccgagaccccaaggcccccaaatggccccgtacacagggcgggggtcagcagttac 2460
Db 2401 gagccccgagaccccaaggcccccaaatggccccgtacacagggcgggggtcagcagttac 2460
QY 2461 gttagtcttggacctggcgccgtggaggtggcgggggggtgcgcgcccagggccttgcgcc 2520
Db 2461 gttagtcttggacctggcgccgtggaggtggcgggggggtgcgcgcccagggccttgcgcc 2520
QY 2521 ttcttggaaacggcttctccccaattgtctcagcgctactcgtataggttaccgagctctctc 2580
Db 2521 ttcttggaaacggcttctccccaattgtctcagcgctactcgtataggttaccgagctctctc 2580
QY 2581 aacaactctagtagaggttctctctctctatatgtataa taaggtatgcactat 2640
Db 2581 aacaactctagtagaggttctctctctctatatgtataa taaggtatgcactat 2640
QY 2641 tcaaataggagcattagctatgtttgttaatgtcactttatgtttatgtgggtaagtcacc 2700

Db 2641 tcaaataggagcattagctatgtttgttaatgtcactttatgtttatgtgggtaagtcacc 2700
QY 2701 taagacactccacgtcacctacgtgtgtctcttaccggctttaataaattcttctgcct 2760
Db 2701 taagacactccacgtcacctacgtgtgtctcttaccggctttaataaattcttctgcct 2760
QY 2761 tgttccatatttactaataattcccttcttccactaaaagaaaaattgttcatcattaagta 2820
Db 2761 tgttccatatttactaataattcccttcttccactaaaagaaaaattgttcatcattaagta 2820
QY 2821 tagtctttagaacatatgaggtcttitaattgggtagggtttttacaaaattacactaataaa 2880
Db 2821 tagtctttagaacatatgaggtcttitaattgggtagggtttttacaaaattacactaataaa 2880
QY 2881 atgtcatataaa tccacgtggttaaacaaatgcagaaaa tgcgactcgtctata ttggaccga 2940
Db 2881 atgtcatataaa tccacgtggttaaacaaatgcagaaaa tgcgactcgtctata ttggaccga 2940
QY 2941 cagttgctattataata tggggccacca tagtagactgacaaataaaattcacctgacaa 3000
Db 2941 cagttgctattataata tggggccacca tagtagactgacaaataaaattcacctgacaa 3000
QY 3001 tctgttctactaataaacacacacaaaaaggagtgcattttccagggcatttttgaata 3060
Db 3001 tctgttctactaataaacacacacaaaaaggagtgcattttccagggcatttttgaata 3060
QY 3061 aaaaacagttaaaaggagtgcaatagaaaaatagggggtgtggaatatgattttgagca 3120
Db 3061 aaaaacagttaaaaggagtgcaatagaaaaatagggggtgtggaatatgattttgagca 3120
QY 3121 cgtcttgaagcgaat tcaactgcccgtcg ttttacaacg tctgtgactgggaaaaacctggc 3180
Db 3121 cgtcttgaagcgaat tcaactgcccgtcg ttttacaacg tctgtgactgggaaaaacctggc 3180
QY 3181 gttaccacactaaatgcgcttcagacacatcccccttccgagctggcgtaataagcga 3240
Db 3181 gttaccacactaaatgcgcttcagacacatcccccttccgagctggcgtaataagcga 3240
QY 3241 gagggccgcgacgatacgcccttcccaacagttgcgcagcctgaattggcgaa tggcgcccg 3300
Db 3241 gagggccgcgacgatacgcccttcccaacagttgcgcagcctgaattggcgaa tggcgcccg 3300
QY 3301 atgcggattttctccttctacgcatctgtcggttatttcaacacgcatatgggtgcatctc 3360
Db 3301 atgcggattttctccttctacgcatctgtcggttatttcaacacgcatatgggtgcatctc 3360
QY 3361 agtacaactctctctgatgcgcagatgttaagccagcccgcacacccgcacacccgct 3420
Db 3361 agtacaactctctctgatgcgcagatgttaagccagcccgcacacccgcacacccgct 3420
QY 3421 gagcgccctgacgggttctgtctcccggcacccgttcacagacaagctgtgacccgtc 3480
Db 3421 gagcgccctgacgggttctgtctcccggcacccgttcacagacaagctgtgacccgtc 3480
QY 3481 tccgggagctgatgtgtcagagggtttcacgcgtaacacccgaaacgcgcgagagaaag 3540
Db 3481 tccgggagctgatgtgtcagagggtttcacgcgtaacacccgaaacgcgcgagagaaag 3540
QY 3541 ggcctcgtgatcgcgctattttttaggttaatgtcatgataataatggttcttagacg 3600
Db 3541 ggcctcgtgatcgcgctattttttaggttaatgtcatgataataatggttcttagacg 3600
QY 3601 tcaagtggcacttttcggggaaaatgtgcgcggaacccctattttgttttttcttaataa 3660
Db 3601 tcaagtggcacttttcggggaaaatgtgcgcggaacccctattttgttttttcttaataa 3660
QY 3661 catcaaatatgataccgctcatgagacaaataacccctgataaattgcctcaataattga 3720
Db 3661 catcaaatatgataccgctcatgagacaaataacccctgataaattgcctcaataattga 3720
QY 3721 aaaaagagagatagattcaacatttccggtgcgccttatcccttttttgggca 3780

Db 3721 aaaaagaagatgagttatcaaatcttcgtgctgccttattcccttttttgcgga 3780
QY 3781 ttgtctctctgtttttgtctaccagaaacgctggtgaaagtataaagatgctgaagat 3840
Db 3781 ttgtctctctgtttttgtctaccagaaacgctggtgaaagtataaagatgctgaagat 3840
QY 3841 cagttgggtgcacagtggtttacatcgaaactggatctcacaacgaggttaagatctcttgag 3900
Db 3841 cagttgggtgcacagtggtttacatcgaaactggatctcacaacgaggttaagatctcttgag 3900
QY 3901 agttttcccccgaagaacggttttcaaatgatgagcaacttttaaaagtctgctatgtggc 3960
Db 3901 agttttcccccgaagaacggttttcaaatgatgagcaacttttaaaagtctgctatgtggc 3960
QY 3961 gcggtattatcccgattgacgcggcggaagagcaactggtcgcgcatatactattct 4020
Db 3961 gcggtattatcccgattgacgcggcggaagagcaactggtcgcgcatatactattct 4020
QY 4021 cagaatgacttgggtgagttactcaccagtcacagaaagacatcttaccggtggcatgaca 4080
Db 4021 cagaatgacttgggtgagttactcaccagtcacagaaagacatcttaccggtggcatgaca 4080
QY 4081 gtaagagaattatgcagtgctgcataaacatgagtatacaactgcggccaacttactt 4140
Db 4081 gtaagagaattatgcagtgctgcataaacatgagtatacaactgcggccaacttactt 4140
QY 4141 ctgacaacgatcggaagacgaaggagctaacccgtttttgacacaacatgggggatcat 4200
Db 4141 ctgacaacgatcggaagacgaaggagctaacccgtttttgacacaacatgggggatcat 4200
QY 4201 gtaactcgcttgatcgttggaacacggagctgaatgaagaccataccaacacgacgagcgt 4260
Db 4201 gtaactcgcttgatcgttggaacacggagctgaatgaagaccataccaacacgacgagcgt 4260
QY 4261 gacacacgatgctgtagcaatggcaacacgcttgcgcaaacattataactggcgaaacta 4320
Db 4261 gacacacgatgctgtagcaatggcaacacgcttgcgcaaacattataactggcgaaacta 4320
QY 4321 cttaactctagcttcccggaacaaatataagactggtgagcgcgataaaagtgcagga 4380
Db 4321 cttaactctagcttcccggaacaaatataagactggtgagcgcgataaaagtgcagga 4380
QY 4381 ccaactctcgctcggcccttcggctgctggtttatgtgataaaatctggagccggt 4440
Db 4381 ccaactctcgctcggcccttcggctgctggtttatgtgataaaatctggagccggt 4440
QY 4441 gagcgtgggtctcggtgatactatgcagcaactgggcccagatggttaagccctccggtatc 4500
Db 4441 gagcgtgggtctcggtgatactatgcagcaactgggcccagatggttaagccctccggtatc 4500
QY 4501 gtagttatctacacgacgagggagtcaggcaactatggtgaaacgaatagacagatcgct 4560
Db 4501 gtagttatctacacgacgagggagtcaggcaactatggtgaaacgaatagacagatcgct 4560
QY 4561 gagatagggtcctcaactgattgaagcattggttaactgtcagaccaaagttaactatata 4620
Db 4561 gagatagggtcctcaactgattgaagcattggttaactgtcagaccaaagttaactatata 4620
QY 4621 ctttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccctttt 4680
Db 4621 ctttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccctttt 4680
QY 4681 gataatctcatgacaaaatcccttaacgtgagttttcttcactgagcgtcagacccc 4740
Db 4681 gataatctcatgacaaaatcccttaacgtgagttttcttcactgagcgtcagacccc 4740
QY 4741 gtgaaagaatcaaaagatctcttgagatccctttttttctgcgcgtataactctgcttg 4800
Db 4741 gtgaaagaatcaaaagatctcttgagatccctttttttctgcgcgtataactctgcttg 4800
QY 4801 caaacaacaaaacacccgctaccagcgggtggtttgtttgcccggatcaagagctaccaact 4860
Db 4801 caaacaacaaaacacccgctaccagcgggtggtttgtttgcccggatcaagagctaccaact 4860

QY 4861 cttttccgaagtaactggcttcagcagagcgagataccaataactgtcccttctagt 4920
Db 4861 cttttccgaagtaactggcttcagcagagcgagataccaataactgtcccttctagt 4920
QY 4921 tagccgtagttagggcaccaacttcaagaactctgtagcacgcctacatacctcgtctg 4980
Db 4921 tagccgtagttagggcaccaacttcaagaactctgtagcacgcctacatacctcgtctg 4980
QY 4981 ctaatcctgttacccgtggtcgtccagtgccagtgccagtgccagtgccagtgccagtgccag 5040
Db 4981 ctaatcctgttacccgtggtcgtccagtgccagtgccagtgccagtgccagtgccagtgccag 5040
QY 5041 tcaagacgatagtataccggataaggcgagcgtcggtcggtcggtcggtcggtcggtcggtcggtcg 5100
Db 5041 tcaagacgatagtataccggataaggcgagcgtcggtcggtcggtcggtcggtcggtcggtcggtcg 5100
QY 5101 cagccacgtctggagcgaaacgactacacacgaaactagatatacctacagcgtgagctatga 5160
Db 5101 cagccacgtctggagcgaaacgactacacacgaaactagatatacctacagcgtgagctatga 5160
QY 5161 gaaagcgcacgcttcccgaagggagaaagcgagcaggtatccggtaaagcgcaggggtc 5220
Db 5161 gaaagcgcacgcttcccgaagggagaaagcgagcaggtatccggtaaagcgcaggggtc 5220
QY 5221 ggaacaggagagcgacgagggagcttccaggggggaaacgcttctttatagtcct 5280
Db 5221 ggaacaggagagcgacgagggagcttccaggggggaaacgcttctttatagtcct 5280
QY 5281 gtcgggtttccacacctctgaactgagcgtcgatgtttgtgagcgtcgtcagggggcg 5340
Db 5281 gtcgggtttccacacctctgaactgagcgtcgatgtttgtgagcgtcgtcagggggcg 5340
QY 5341 agcctatggaaaaacgacgacgagcgttctttacggttctcgtggttctgtggtcct 5400
Db 5341 agcctatggaaaaacgacgacgagcgttctttacggttctcgtggttctgtggtcct 5400
QY 5401 ttgtctacatgtttcttctcgttattccctgattctctgtgataacccgtattaccgcc 5460
Db 5401 ttgtctacatgtttcttctcgttattccctgattctctgtgataacccgtattaccgcc 5460
QY 5461 ttgtgagtgagctataccgctccgcagcagcagcagcagcagcagcagcagcagcagcagcagc 5520
Db 5461 ttgtgagtgagctataccgctccgcagcagcagcagcagcagcagcagcagcagcagcagcagc 5520
QY 5521 gaggaagcggaagcgcccaatacgaacacccctctcccggttggccgattcat 5580
Db 5521 gaggaagcggaagcgcccaatacgaacacccctctcccggttggccgattcat 5580
QY 5581 taatgcagctggcagcagaggtttcccgactggaaagcgggcagtgagcgaacgcaatt 5640
Db 5581 taatgcagctggcagcagaggtttcccgactggaaagcgggcagtgagcgaacgcaatt 5640
QY 5641 aatgtagttagctactcatttagccacccagcgttaccatttatgtctcggctcgt 5700
Db 5641 aatgtagttagctactcatttagccacccagcgttaccatttatgtctcggctcgt 5700
QY 5701 atgtgtgtggaattgtgagcggataacaattttcacacaggaacacgctatgacccatg 5760
Db 5701 atgtgtgtggaattgtgagcggataacaattttcacacaggaacacgctatgacccatg 5760
QY 5761 tacgcca 5767
Db 5761 tacgcca 5767

RESULT 2

AAV63741

ID AAV63741 standard; DNA; 5897 BP.

XX

AAV63741;

XX

DT 12-APR-1999 (first entry)

Db 627 gagagaaacaggggagctctagaggatccagctgaaggtctcgaaaggcagtgccacggag 686
QY 744 atcaagcattctacttctattgag-----caatttaaatcatcttctttaagca 794
Db 687 gaagctgatatgttgagcaagctgtgagtagagcaaacctctccataataaccagca 746
QY 795 aaagcaattttctgaaattttcaccaatttaacaaagatagccatggctccccgcagtg 854
Db 747 ccaccaagtcagggaatccccagatcaagtgcaaaaggtccgctgtttctctctctgc 806
QY 855 tctgctgcacacgcct-----tccctgg 877
Db 807 tctgctgctgactaaactctgggttatgattcgtttgagtaattttggggaaagctctcttg 866
QY 878 ctctcccaactctctctctctctctctggctctgggtgagagagtggggctgagggcc 937
Db 867 ctgctccacacatgctccattcgaattaccgtgttagcaaggcgcaaaagttagc 926
QY 938 gggagatgcaga-----gctgctggtagcgtgctgctggggccgctg 981
Db 927 ttgattgattgattgactatgattgctttctggaccctgacgctgacgtgagcgatc 986
QY 982 cgggcatctgcctgaagaccccggggcccctgtctctgctttcc-----tgg 1030
Db 987 tggggccatttctccaggcacgggataagcatctcagccaatggctcctctgtagaacc 1046
QY 1031 gcatccccctttggggacacccatggaccctgctgctttctgcacccgagcccaagc 1090
Db 1047 ccaaccggaataatcaaaaaacacgcagcctgtggcattcagctggtgacgcgaaac 1106
QY 1091 agccttggctcaggggtgtagacgtctaacacctccagagt-----tctgcta 1139
Db 1107 tgtggaattgacagcgttgtgggaaagcgcgttacagaagaagccgggcaattctctgt 1166
QY 1140 ccaatatggacacctataccaggttttgaggccacgcagatgtgaaccccaaccg 1199
Db 1167 ccaggcagttttaaactcagttctgcgagtgccagatattcgttaattatgcgggcaacgctc 1226
QY 1200 tgagctgagcagagactgctgtacctcaactgtgtgacaccataccccggcctacatc 1259
Db 1227 tggatcagcgcgaagctctttaaccgaaaggttggggcagcgtatcgtgctgct 1286
QY 1260 cccacccctgtcctgctgctgctatgaggggtgcttctacagtggggctcctcctct 1319
Db 1287 ttcgattgctgctactattacggcaagtggtggtcaataatcaggaagtgtggagcat 1346
QY 1320 ggacgtgtacgatggcctctctgtgtacagccgcagagactgtgctgtgctccatgaa 1379
Db 1347 -caggcggtatcacccatttgaagccgatgtcacgcgctatgttattgcccgggaaaa 1405
QY 1380 ctaccgggtgggagcctttggcttctggtccctggcggggagcgcagagggcccgccgcaa 1439
Db 1406 tgcagtatcacctgttgttgaaacacgaactgaactggcagactatcccgcggaat 1465
QY 1440 tgggttctctctggatcaagggctggccctgcaagtggg----- 1477
Db 1466 ggtgattacgcagaaacggcgaagaaagcagttctactccatgatttctttaacta 1525
QY 1478 tgcaggaacagtggcag-----ccttgggggtaccc 1511
Db 1526 tgcggaatccaatcgacgcgtaattgctctacacacgcgcgaacacctggggagcagatat 1585
QY 1512 gacatcagtgacgctgtttgggagagcgcgggagccgctcgtggtggcatgcaacctgct 1571
Db 1586 caccgtggtagcagcatgtcgcgaagactgttaaccagcgtctgtgactggcagtggt 1645
QY 1572 gtccccccagccgggctgtttcacaggggcgtgtcgtgacagcgggtgcccccaatgg 1631
Db 1646 ggcacaatggtgatgtcagcgttgaactgctgtatgctggatcaacaggtgtgtgcaactgg 1705
QY 1632 accctgggcacagtggtggcatggagagggcccgctgcagggccacgcagctggccacct 1691

Db 1706 acaaggcactagcgggactttgcaagtgggtgaatcgcgcaacctctggcaacgggtgaagg 1765
QY 1692 tgtgggctgtccctccagggcgccactggtgggaaatgacacagagctggtagcctgc---ct 1748
Db 1766 ttatctctatgaactgtgctcacagcaaaagccagacagagtgtgatatctaccgct 1825
QY 1749 tcggacacgaccagcgcaggctcctgtgtgaacacgaatggcacgctgctgctcaagaag 1808
Db 1826 tcgctgctggcatccggtcagtgccagtggaaggcgcaacagttcctgattaaaccacaacc 1885
QY 1809 cgtctccggttctctctgctgtgctggtgtagatgagagacttctcagtgacacccacga 1868
Db 1886 gtctactttactggcttggctgctgaatgagctgagacttgcgtggcaaaagattcga 1945
QY 1869 ggccctcatcaacgcgggagacttccacggcctgcagcgtgctggtggtggtggtggaagg 1928
Db 1946 taacgtgctgattggtgcacgaccacgcattaaatggactggatggggccaaactcctaccg 2005
QY 1929 tgagggctcgtattttctggttttacggggcccaaggcttcagcaaaagacacagtgctct 1988
Db 2006 taactcgcattaccccttacgctgaagagatgctcgcactggcgagatgaacatggcatcgt 2065
QY 1989 catcagccggcgagttcctgcccgggtgcgggtcggggttcccccagtgaaagtacct 2048
Db 2066 ggtgattgataaactgctgctgcggcttaacctctctttagggcattgtgtttcgaagc 2125
QY 2049 ggagcgcgagctgtgctcctgattacacagactggctgcatccccgagagaccggcgacg 2108
Db 2126 ggcaacaacgcgaagaactgttacagcgaagagcgactcaacgggggaaactcagcaagc 2185
QY 2109 cctgaggggagccctcagcagatgtg---gtggcgacaccaaatgtcgtgtgcgccgtggc 2165
Db 2186 gccttacaggcgtataaagagctgtagcgcgtgacaaaaacacccaagcgtgggtgat 2245
QY 2166 ccagctggctggcgactgctgcccaggtgcccgggtctacgctcactctttgaaca 2225
Db 2246 gtggagtattgcacaagaccgatacccgctccgcaagtgcagcgggaattatctgcacct 2305
QY 2226 cctgtcttcacgctctcctggccctgtggtgggtgggtgcccacagcgttacagatcga 2285
Db 2306 ggcggaagcaacgcgtaaaactgacccgacgcgctccgcatcactgcgtcaatgtatgtt 2365
QY 2286 gtcaatctttgggatac-----ccctggagccctctcgaaaactcacaggca 2331
Db 2366 ctgcgacgctcacaccgataccatcagcagatctcttga tgtgctgtgcctgaaccgtta 2425
QY 2332 gaggagaaaaatcttcgccagcagctgatacgtactggtggccaaacttgcgccacaggg 2391
Db 2426 ttacgcatggtatgtccaaagcggcgtatttgaaacgcgcagagaaggctactggaaaaa 2485
QY 2392 gatcccaatgagccccag--accccaaggcccccaaatggcccccgctacacggcggg- 2448
Db 2486 actctggcctggcaggagaaactgcctcagcgcgtattcatcaccgaaatcacggcgtgga 2545
QY 2449 -----gctcagcagctacgttagctggaacctggcccgctggaggtgcgcgggg 2498
Db 2546 tacgttagcgggctgcactcaatgtacacccgacatgtgagtgaaagagtatcagtgtc 2605
QY 2499 gctcgcgccacagcctgcgcttctggaaacgcttctccccaatttgcctgcgcgtac 2558
Db 2606 atggctggatatatacacgcgctcttctgctcgcgcagcgcgcgtcgcggtaaacaggt 2665
QY 2559 ctgatagggtaccgagctctctcaacaatctagtagagttgtctcctatataatgtcaat 2618
Db 2666 atggaaatttcgcgattttgcacctgcgaagcatattgcggttggcggttaacaagaa 2725
QY 2619 aaggtatgctgatatgcactattcaaataggagcattagctatgtttgttaagtcaact 2678
Db 2726 agggattctcactcgcgacccgcaaacccgaagtcggcgctttctctgctgcaaaaacgctg 2785
QY 2679 tatgttatgtgggttaagtcacccctaaagacactccacgtacactagttgtgtctctaccg 2738
Db 2786 gactggcatgaactcgtggtgaaaaacccgcagcagggggcaaacatgaatcaacaactc 2845

Db 4188 ccttgacgggctgtctgctccggcatccgcttcacagacaagctgtgaccgtctccggg 4247
QY 3487 agtgcattgttcagagggttttcacgctcatcacccgaacgcgcgcagacgaagggcctc 3546
Db 4248 agtgcattgttcagagggttttcacgctcatcacccgaacgcgcgcagacgaagggcctc 4307
QY 3547 ggtacgcctctttttatagatttaagtcatgataaataatggtttcttagacgtcagct 3606
Db 4308 ggtacgcctctttttatagatttaagtcatgataaataatggtttcttagacgtcagct 4367
QY 3607 ggcacttttcggggaatgtgcgcgaacccctattgttttattttctaaatacatcca 3666
Db 4368 ggcacttttcggggaatgtgcgcgaacccctattgttttattttctaaatacatcca 4427
QY 3667 aatatgtatccgctcatgagacaataaacctgtataaattgtctcaataatattgaaaaagg 3726
Db 4428 aatatgtatccgctcatgagacaataaacctgtataaattgtctcaataatattgaaaaagg 4487
QY 3727 aagatgatgattcaaacattccgtgtcgcgcctattcccttttttgcgcattttgc 3786
Db 4488 aagatgatgattcaaacattccgtgtcgcgcctattcccttttttgcgcattttgc 4547
QY 3787 ctctcgtttttgctcacccagaacgcgtgtgaaagtataaagatgctgaagatcagttg 3846
Db 4548 ctctcgtttttgctcacccagaacgcgtgtgaaagtataaagatgctgaagatcagttg 4607
QY 3847 ggtcacgagtggtttacatcgaaactggatctcaacagcggtaagatcccttgagagtttt 3906
Db 4608 ggtcacgagtggtttacatcgaaactggatctcaacagcggtaagatcccttgagagtttt 4667
QY 3907 cgcgccgaagaacggttttccaatgatgagacatttttaaagtctcgtatgtgcgcggtta 3966
Db 4668 cgcgccgaagaacggttttccaatgatgagacatttttaaagtctcgtatgtgcgcggtta 4727
QY 3967 ttatcccgattgacgcgcgggcaagagcaactcggtcgcgcgcataacactattctcagaat 4026
Db 4728 ttatcccgattgacgcgcgggcaagagcaactcggtcgcgcgcataacactattctcagaat 4787
QY 4027 gacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga 4086
Db 4788 gacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga 4847
QY 4087 gaattatgagtgctgccataaccatgattgataacactgcgcgcacacttaactttctgaca 4146
Db 4848 gaattatgagtgctgccataaccatgattgataacactgcgcgcacacttaactttctgaca 4907
QY 4147 acgatacgaggaccgaagagctaacccgttttttgacaaacatgggggacatcatgtaact 4206
Db 4908 acgatacgaggaccgaagagctaacccgttttttgacaaacatgggggacatcatgtaact 4967
QY 4207 cgccttgatcgttgggaaccggagctgaatgaagccataaccaaaacgagcgtgacacc 4266
Db 4968 cgccttgatcgttgggaaccggagctgaatgaagccataaccaaaacgagcgtgacacc 5027
QY 4267 acgatacgcttagcaatggcaacaacggttgcgaacactattaaactggcggaacttaact 4326
Db 5028 acgatacgcttagcaatggcaacaacggttgcgaacactattaaactggcggaacttaact 5087
QY 4327 ctagcttcccggaacaataatagactggatggaggcggaataaagtgtcaggaaccaact 4386
Db 5088 ctagcttcccggaacaataatagactggatggaggcggaataaagtgtcaggaaccaact 5147
QY 4387 ctgcgctcggccctccggctggctgggttataatgtgataaaatctggagccggtgagcgt 4446
Db 5148 ctgcgctcggccctccggctggctgggttataatgtgataaaatctggagccggtgagcgt 5207
QY 4447 gggctcgcggttatcatgacacactggggccagatggttaagccctcccgatcgtagtt 4506
Db 5208 gggctcgcggttatcatgacacactggggccagatggttaagccctcccgatcgtagtt 5267
QY 4507 actacacgacgggagtcaggcaactatggatgaacgaataagacagatcgcgtgagata 4566
Db 4566 actacacgacgggagtcaggcaactatggatgaacgaataagacagatcgcgtgagata

Db 5268 atctacacgacggggagtcaggcaactatggatgaacgaataagacagatcgcgtgagata 5327
QY 4567 ggtgcctcactgattaaagcatgtgtaactgtcagacaagttttactcatataacttttag 4626
Db 5328 ggtgcctcactgattaaagcatgtgtaactgtcagacaagttttactcatataacttttag 5387
QY 4627 attgatttaaaacttcaatttttaatttaaaagagatctagggtgaagatcctttttgataat 4686
Db 5388 attgatttaaaacttcaatttttaatttaaaagagatctagggtgaagatcctttttgataat 5447
QY 4687 ctcatgaccaaataccttaacgtgagttttcgttccactgagcgtcagaccccgtagaa 4746
Db 5448 ctcatgaccaaataccttaacgtgagttttcgttccactgagcgtcagaccccgtagaa 5507
QY 4747 aagatacaaggagatcttcttgagatccttttttctgcgcgtaactcgtcgtttgcaaca 4806
Db 5508 aagatacaaggagatcttcttgagatccttttttctgcgcgtaactcgtcgtttgcaaca 5567
QY 4807 aaaaaaacccgctaccacgctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4866
Db 5568 aaaaaaacccgctaccacgctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5627
QY 4867 ccgaaggtaactggcttcagcagagcgcagataccaaaataactgtccttctagtgtagccg 4926
Db 5628 ccgaaggtaactggcttcagcagagcgcagataccaaaataactgtccttctagtgtagccg 5687
QY 4927 tagttaggccacacacttcaagaactctgtagcaccgcctacatacctcgtcctgtctaactc 4986
Db 5688 tagttaggccacacacttcaagaactctgtagcaccgcctacatacctcgtcctgtctaactc 5747
QY 4987 ctgttaccagtggtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5046
Db 5748 ctgttaccagtggtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5807
QY 5047 cgatagttaccggataaagcgcgcgcgtcgcgtcgcgtcgcgtcgcgtcgcgtcgcgtcgcgt 5106
Db 5808 cgatagttaccggataaagcgcgcgcgtcgcgtcgcgtcgcgtcgcgtcgcgtcgcgtcgcgt 5867
QY 5107 agcttggagcgaaacacactacaccgaactgagatacactacagcgtgagctatgagaaagc 5166
Db 5868 agcttggagcgaaacacactacaccgaactgagatacactacagcgtgagctatgagaaagc 5927
QY 5167 gccacgcttccgaagggagaaagcgcgcaggtatccggtaagcggcgaggggtcgggaaca 5226
Db 5928 gccacgcttccgaagggagaaagcgcgcaggtatccggtaagcggcgaggggtcgggaaca 5987
QY 5227 ggagagcgacagggaggttccagggggaaacgcgcgtggtatcttatagtcgtcggg 5286
Db 5988 ggagagcgacagggaggttccagggggaaacgcgcgtggtatcttatagtcgtcggg 6047
QY 5287 ttctgcacactctgacttgagcgtcgatttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5346
Db 6048 ttctgcacactctgacttgagcgtcgatttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 6107
QY 5347 tggaaaaacgcagcaacgcgcctttttacgggttccgtggccttttgcgtggccttttgcgt 5406
Db 6108 tggaaaaacgcagcaacgcgcctttttacgggttccgtggccttttgcgtggccttttgcgt 6167
QY 5407 cacatgttttctcgtcgttatccctcgtattctgtggataaacctgattaccgccttttgag 5466
Db 6168 cacatgttttctcgtcgttatccctcgtattctgtggataaacctgattaccgccttttgag 6227
QY 5467 tgagctgataccgcctcgcgcagccgaacacgcagcgcgcagcgcagcgcagcgcagcgcagcgcag 5526
Db 6228 tgagctgataccgcctcgcgcagccgaacacgcagcgcgcagcgcagcgcagcgcagcgcagcgcag 6287
QY 5527 gcggaagacgcaccaataacgcaaacgcctctcccgcgcttcccgcgcttcccgcgcttcccgcgct 5586
Db 6288 gcggaagacgcaccaataacgcaaacgcctctcccgcgcttcccgcgcttcccgcgcttcccgcgct 6347
QY 5587 agctggcacgacaggttttcccgactggaaagcgcgcagtgagcgcgaacgaacgaacgaacgaacga 5646
Db 6348 agctggcacgacaggttttcccgactggaaagcgcgcagtgagcgcgaacgaacgaacgaacgaacga 6407

QY 5647 agtagctactcattaggcaacccagcgtttacacatttatgtcttcggtcgtatgttg 5706
|||||
Db 6408 agtagctactcattaggcaacccagcgtttacacatttatgtcttcggtcgtatgttg 6467

QY 5707 tgtggaattgtgagcgataacaaatttcacacaggaacagctatgacctattacgcc 5766
|||||
Db 6468 tgtggaattgtgagcgataacaaatttcacacaggaacagctatgacctattacgcc 6527

QY 5767 a 5767
|
Db 6528 a 6528

RESULT 4
AAD04947
ID AAD04947 standard; DNA; 3858 BP.
XX
AC AAD04947;
DT 17-JUL-2001 (first entry)
XX
DE Plasmid pRK50 used to test Cre recombinase mediated inversion.
XX
KW Gene trapping construct; conditional mutation; unidirectional inversion;
XX recombinase recognition sequence; RRS; disruption cassette;
KW selection cassette; transgenic organism; pRK50; Cre recombinase; ds.
XX
OS Chimeric - Adenovirus.
OS Chimeric - Unidentified.
OS Chimeric - ECMV virus.
XX
PN WO200129208-A1.
XX
XX 26-APR-2001.
PD
XX
PF 16-OCT-2000; 2000WO-EP10162.
XX
XX 16-OCT-1999; 99EP-0120592.
PR
XX 27-OCT-1999; 99US-0162016.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX
PI Kuehn R, Von Melchener H, Altschmied J;
XX
XX WPI; 2001-308486/32.
DR
XX
XX New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX
PS Example 3; Page 73-74; 78pp; English.
XX
XX The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is pRK50 vector, which is used to test Cre
CC recombinase mediated inversion, which is related to the invention.
XX
SQ Sequence 3858 BP; 943 A; 959 C; 994 G; 962 T; 0 other;

Query Match 45.7%; Score 2637; DB 22; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaactggcgtctgttttacacacgtctgactgggaaaaacccctggcgttacccaac 3190
|||||
Db 1187 cgaattcaactggcgtctgttttacacacgtctgactgggaaaaacccctggcgttacccaac 1246

QY 3191 ttaatcgcttgacgacacatccccccttcgcagctgagcgaatagcgaagagcccgca 3250
|||||
Db 1247 ttaatcgcttgacgacacatccccccttcgcagctgagcgaatagcgaagagcccgca 1306

QY 3251 ccgatacgcccttcccaacagttgcgacgctgaatggcgaatggcgcctgatcggtatt 3310
|||||
Db 1307 ccgatacgcccttcccaacagttgcgacgctgaatggcgaatggcgcctgatcggtatt 1366

QY 3311 ttctccttagcaatctgtgcggtattttcacacgcgcataatggtgcactctcagtcacaaatct 3370
|||||
Db 1367 ttctccttagcaatctgtgcggtattttcacacgcgcataatggtgcactctcagtcacaaatct 1426

QY 3371 gctctgatgcgcatagttaagccagccgcacacccgcacacccgcgtgacgcgcct 3430
|||||
Db 1427 gctctgatgcgcatagttaagccagccgcacacccgcacacccgcgtgacgcgcct 1486

QY 3431 gacgggcttctgtctcccgccatccgcttcacagacaaagctgtgacgctctccggagct 3490
|||||
Db 1487 gacgggcttctgtctcccgccatccgcttcacagacaaagctgtgacgctctccggagct 1546

QY 3491 gcatgtctcagaggttttcccgctcatcacccgaacccgcgagacgaagggcctcgtga 3550
|||||
Db 1547 gcatgtctcagaggttttcccgctcatcacccgaacccgcgagacgaagggcctcgtga 1606

QY 3551 tacgctattttataggttaattgataataatggtttctcttagacgtcaggttgca 3610
|||||
Db 1607 tacgctattttataggttaattgataataatggtttctcttagacgtcaggttgca 1666

QY 3611 ctttcggggaaatgtgcgcggaacccctattgtttattttttaaatacattcaata 3670
|||||
Db 1667 ctttcggggaaatgtgcgcggaacccctattgtttattttttaaatacattcaata 1726

QY 3671 tgtatccgctcatgagacataacccctgataaaatgcttcataataattgaaaaaggaaga 3730
|||||
Db 1727 tgtatccgctcatgagacataacccctgataaaatgcttcataataattgaaaaaggaaga 1786

QY 3731 gtatgatttcaacatttcgctgcgccttattcccttttttggcgcaatttgccttc 3790
|||||
Db 1787 gtatgatttcaacatttcgctgcgccttattcccttttttggcgcaatttgccttc 1846

QY 3791 ctgtttttgtcccccagaacacgtggtgaaagttaaagatgctgaagatacagttgggtg 3850
|||||
Db 1847 ctgtttttgtcccccagaacacgtggtgaaagttaaagatgctgaagatacagttgggtg 1906

QY 3851 cagcagtggtttacatcgaactggtatcacaacgcggtgaagatcccttgagagtttccgc 3910
|||||
Db 1907 cagcagtggtttacatcgaactggtatcacaacgcggtgaagatcccttgagagtttccgc 1966

QY 3911 ccgaagaacgtttttccaatgatgagcaacttttaaagtcttgcctatgtggcggtattat 3970
|||||
Db 1967 ccgaagaacgtttttccaatgatgagcaacttttaaagtcttgcctatgtggcggtattat 2026

QY 3971 ccgctattgacgcggggaagacaaactcgtgcgcgcatacactattctcagaatgact 4030
|||||
Db 2027 ccgctattgacgcggggaagacaaactcgtgcgcgcatacactattctcagaatgact 2086

QY 4031 tgtttgagtactaccagtcacagaaaagcatcttcaggatggcagatgacagtaagaat 4090
|||||
Db 2087 tgtttgagtactaccagtcacagaaaagcatcttcaggatggcagatgacagtaagaat 2146

QY 4091 tatgcagtgctgcataaacatgagtataacactcgcgcgcacacttactctgacacaga 4150
|||||
Db 2147 tatgcagtgctgcataaacatgagtataacactcgcgcgcacacttactctgacacaga 2206

[illegible]

RESULT	5	
AAAD09269		
ID	AAAD09269	standard; DNA; 4960 BP.
XX		
AC	AAAD09269;	
XX		
DT	12-SEP-2001	(first entry)
XX		
DE	pcMV-I-Cre-pA vector DNA.	
XX		
KW	DNA recombinase domain; protein transduction domain; PTD;	
KW	gene alteration; fusion protein; Human immunodeficiency virus;	
KW	HIV; pcMV-I-Cre-pA vector; ds.	
XX		
OS	Unidentified.	
XX		
PN	WO200149832-A2.	
XX		
PD	12-JUL-2001.	
XX		
PF	05-JAN-2001; 2001WO-EP00060.	
XX		
PR	07-JAN-2000; 2000EP-0100351.	
PR	10-NOV-2000; 2000EP-0124595.	
XX		
PA	(ARTE-) ARTEMIS PHARM GMBH.	
XX		
PI	Schwenk F;	
XX		
DR	WPI; 2001-441873/47.	
XX		
PT	Using site-specific DNA recombinase domain/protein transduction domain	
PT	fusion proteins for inducing target gene alterations in organisms or	
PT	cell cultures -	
XX		
PS	Example; Page 62-63; 85pp; English.	
XX		

CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAR peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a pCMV-I-Cre-pA vector DNA.
XX
SQ

Sequence 4960 BP; 1225 A; 1213 C; 1296 G; 1226 T; 0 other;

Query Match 45.7%; Score 2637; DB 22; Length 4960;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3131 cgaattcactgcccgcgttttacacgctgtagctgggaaacccctggcttacccaac 3190
DB 2289 cgaattcactgcccgcgttttacacgctgtagctgggaaacccctggcttacccaac 2348
QY 3191 ttaatcgcccttcagacacatccccccttcgcagctggcgtaatacggaagggcccgca 3250
DB 2349 ttaatcgcccttcagacacatccccccttcgcagctggcgtaatacggaagggcccgca 2408
QY 3251 ccgagtcgcccctcccaaacagttgcgagcctgaatggcgaatggcgcgtgatt 3310
DB 2409 ccgagtcgcccctcccaaacagttgcgagcctgaatggcgaatggcgcgtgatt 2468
QY 3311 ttctcttacgcatgtgctggtatttacacgcgcataatggtgcactccagtaaatct 3370
DB 2469 ttctcttacgcatgtgctggtatttacacgcgcataatggtgcactccagtaaatct 2528
QY 3371 gctctgatgcgcacatagtttaagccagcccgacaccccgacaccccgctgacgcgcct 3430
DB 2529 gctctgatgcgcacatagtttaagccagcccgacaccccgctgacgcgcct 2588
QY 3431 gacgggttgctgtccggcgcacgcgttacagacaaagctgacgcgtccggagct 3490
DB 2589 gacgggttgctgtccggcgcacgcgttacagacaaagctgacgcgtccggagct 2648
QY 3491 gcatgtgcagaggttttcacgcgtcatcacccgaacgcgcgcagagaaaggccctcgtga 3550
DB 2649 gcatgtgcagaggttttcacgcgtcatcacccgaacgcgcgcagagaaaggccctcgtga 2708
QY 3551 tacgcctattttataggttaatgctatgataataatggtttcttagacgctcaggtggca 3610
DB 2709 tacgcctattttataggttaatgctatgataataatggtttcttagacgctcaggtggca 2768
QY 3611 cttttcggggaaatgctgcgggaacccctattgtttatttttcaataacattcaata 3670
DB 2769 cttttcggggaaatgctgcgggaacccctattgtttatttttcaataacattcaata 2828
QY 3671 tgatccgctcatgagacaataaccctgataaattgcttcaataattgaaaggaaga 3730
DB 2829 tgatccgctcatgagacaataaccctgataaattgcttcaataattgaaaggaaga 2888
QY 3731 gtagtgatttcaacatttcctgctgcgccttattcccttttttgcggcaatttgccttc 3790
DB 2889 gtagtgatttcaacatttcctgctgcgccttattcccttttttgcggcaatttgccttc 2948
QY 3791 ctgttttttgctcaaccagaaaacgtggtgaaagtataaagatgctgaagatacagttgggtg 3850
DB 2949 ctgttttttgctcaaccagaaaacgtggtgaaagtataaagatgctgaagatacagttgggtg 3008
QY 3851 caccgagtggtttacatcgaaactggatctcaacacgcgcgttaagatccttgagagtttcgcc 3910
DB 3009 caccgagtggtttacatcgaaactggatctcaacacgcgcgttaagatccttgagagtttcgcc 3068
QY 3911 ccgaagaacggttttccaatgatgagcaactttttaaagttctgtatgtgcccgggtattat 3970
DB 3069 ccgaagaacggttttccaatgatgagcaactttttaaagttctgtatgtgcccgggtattat 3128
QY 3971 ccggtattgacgcccgggcaaggaactcggctccgcgcgcatacactattctcagaatgact 4030
```

```
DB 3129 ccggtattgaccccggaagagcaactcggcgcgcatacactattctcagaatgact 3188
QY 4031 tgggtgactactcaccagtcacagaaaaacatcttacggatggcatgacagtaagagaat 4090
DB 3189 tgggtgactactcaccagtcacagaaaaacatcttacggatggcatgacagtaagagaat 3248
QY 4091 tatgcagtgtcgcataaccatgagtgataaacactgcggcacaacttactcttgacaacga 4150
DB 3249 tatgcagtgtcgcataaccatgagtgataaacactgcggcacaacttactcttgacaacga 3308
QY 4151 tcggaggaccgaagagctaacccgttttttgcacaacatgggggatactgttaactcgc 4210
DB 3309 tcggaggaccgaagagctaacccgttttttgcacaacatgggggatactgttaactcgc 3368
QY 4211 ttgatcgttgggaacgcggagctgaatgaagccataccacaacacgagcgtgacacaga 4270
DB 3369 ttgatcgttgggaacgcggagctgaatgaagccataccacaacacgagcgtgacacaga 3428
QY 4271 tgcctgtagcaatggcaacaacgcttcgcacaactattaaactggcgaacttactcttag 4330
DB 3429 tgcctgtagcaatggcaacaacgcttcgcacaactattaaactggcgaacttactcttag 3488
QY 4331 cttcccggaacaattaatagactggatggagcgggataaaagtgcaggaccactctgc 4390
DB 3489 cttcccggaacaattaatagactggatggagcgggataaaagtgcaggaccactctgc 3548
QY 4391 gtcctgcccctccgcgtggttatttgcgtgataaaactctgagccgctgagcgtgggt 4450
DB 3549 gtcctgcccctccgcgtggttatttgcgtgataaaactctgagccgctgagcgtgggt 3608
QY 4451 ctgcggttatcatctgcagcactggggccagatggtaagccctcccgatcgttagttatct 4510
DB 3609 ctgcggttatcatctgcagcactggggccagatggtaagccctcccgatcgttagttatct 3668
QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgcgtgagatggcg 4570
DB 3669 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgcgtgagatggcg 3728
QY 4571 cctcactgattaaagcattggttaactgcagacccaagtttactcatatatacttttagattg 4630
DB 3729 cctcactgattaaagcattggttaactgcagacccaagtttactcatatatacttttagattg 3788
QY 4631 atttaaaactcatttttaattaaagatctagtggaagcctcttttttgataatctca 4690
DB 3789 atttaaaactcatttttaattaaagatctagtggaagcctcttttttgataatctca 3848
QY 4691 tgacaaaaatcccttaacgctgagtttctgcactgagcgtcagaccccgtagaaaaaga 4750
DB 3849 tgacaaaaatcccttaacgctgagtttctgcactgagcgtcagaccccgtagaaaaaga 3908
QY 4751 tcaaaaggatctcttgagatccctttttctgcgcgttaactcgtcttgcaaacaaaaa 4810
DB 3909 tcaaaaggatctcttgagatccctttttctgcgcgttaactcgtcttgcaaacaaaaa 3968
QY 4811 aaccaccgtaccagcgggtggtttgttgcgggatacaagagctaccaactcttttcoga 4870
DB 3969 aaccaccgtaccagcgggtggtttgttgcgggatacaagagctaccaactcttttcoga 4028
QY 4871 aggttaactggtctcagcagagcgcagataacaaaaactgtccttcttagtgagcgttagt 4930
DB 4029 aggttaactggtctcagcagagcgcagataacaaaaactgtccttcttagtgagcgttagt 4088
QY 4931 taggcacacacttcaagaactctgtagcacccctcacatacctcgtctcgtctaaacctgt 4990
DB 4089 taggcacacacttcaagaactctgtagcacccctcacatacctcgtctcgtctaaacctgt 4148
QY 4991 taccagtggtctgctccagtgccgataaagctgtcttaccgggttgactcaaacgact 5050
DB 4149 taccagtggtctgctccagtgccgataaagctgtcttaccgggttgactcaaacgact 4208
QY 5051 agttaccggataaagcgcagcgttgcggctgaacgggggttcgtgcacacagccagct 5110
```

Db 4209 agttaccggataaagcgacggtcggtggaacgggggttcgtgcacacagccagct 4268
Qy 5111 tggagcgaaacacactacacccaactgagatacctacagctgagctatgagaaagcgcca 5170
Db 4269 tggagcgaaacacactacacccaactgagatacctacagctgagctatgagaaagcgcca 4328
Qy 5171 cgcctccgaaggagaaagcgacagctatccggttaagcgcaaggctcggaacagag 5230
Db 4329 cgcctccgaaggagaaagcgacagctatccggttaagcgcaaggctcggaacagag 4388
Qy 5231 agcgacagaggagcttcacagggggaacgcctggtatctttatagtctcctgcgggttc 5290
Db 4389 agcgacagaggagcttcacagggggaacgcctggtatctttatagtctcctgcgggttc 4448
Qy 5291 gccacctgacttgagcgtcgatattttgtgatgctcgtcagggggcgagcctatgga 5350
Db 4449 gccacctgacttgagcgtcgatattttgtgatgctcgtcagggggcgagcctatgga 4508
Qy 5351 aaagcgccagcaacgcgcctttttacaggttccctggtccttttgccttttgccttc 5410
Db 4509 aaagcgccagcaacgcgcctttttacaggttccctggtccttttgccttttgccttc 4568
Qy 5411 tgtttcttcctgcttaccctgattctgtgataaacgctattaccgccttttgagtga 5470
Db 4569 tgtttcttcctgcttaccctgattctgtgataaacgctattaccgccttttgagtga 4628
Qy 5471 ctgatcagcctgcgcgcgcggaacgcagcagcagcagcagcagcagcagcagcagc 5530
Db 4629 ctgatcagcctgcgcgcgcggaacgcagcagcagcagcagcagcagcagcagcagc 4688
Qy 5531 aagagcgcccaatacagcaaacgcctctcccgcgcttcccgcgcttcccgcgcttcc 5590
Db 4689 aagagcgcccaatacagcaaacgcctctcccgcgcttcccgcgcttcccgcgcttcc 4748
Qy 5591 ggcacgacaggtttcccgactggaagcggcagtgagcgcgaacgaatattgtgagtt 5650
Db 4749 ggcacgacaggtttcccgactggaagcggcagtgagcgcgaacgaatattgtgagtt 4808
Qy 5651 agctcactcattagcagccccaggtttacactttatgcttccgctcgtatggtgtg 5710
Db 4809 agctcactcattagcagccccaggtttacactttatgcttccgctcgtatggtgtg 4868
Qy 5711 gaattgtagcggataaacaatttcacaggaagaaacagctatgacatgattacgcca 5767
Db 4869 gaattgtagcggataaacaatttcacaggaagaaacagctatgacatgattacgcca 4925

RESULT 6

AAD09270
ID AAD09270 standard; DNA: 7332 BP.
XX
AC AAD09270;
XX
DT 12-SEP-2001 (first entry)
XX
DE pCMV-I-beta-pA vector DNA related to the invention.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW gene alteration; fusion protein; Human immunodeficiency virus;
KW HIV; pCMV-I-beta-pA vector; ds.
XX
OS Unidentified.
XX
PN WO200149832-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-EP00060.
XX
PR 07-JAN-2000; 2000EP-0100351.
XX 10-NOV-2000; 2000EP-0124595.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.

XX Schwenk F;
PI
XX
DR WPI; 2001-441873/47.
XX
PT Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PI cell cultures -
XX
PS Example; Page 63-65; 85pp; English.
XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAR peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a pCMV-I-beta-pA vector
CC DNA related to the invention.
XX
SQ Sequence 7332 BP; 1718 A; 1895 C; 1964 G; 1755 T; 0 other;

Query Match 45.7%; Score 2637; DB 22; Length 7332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3131 cgaattcactggtcgctgttttacaacgctgactgggaaacccctggttaccacac 3190
Db 4661 cgaattcactggtcgctgttttacaacgctgactgggaaacccctggttaccacac 4720
Qy 3191 ttaatgccttcgacacacatccccctttccagctggcgtaataagcgaagggccgcga 3250
Db 4721 ttaatgccttcgacacacatccccctttccagctggcgtaataagcgaagggccgcga 4780
Qy 3251 ccgatacgccctcccaacagttgcgagcctgaatggcgaatggcgctgagcggtatt 3310
Db 4781 ccgatacgccctcccaacagttgcgagcctgaatggcgaatggcgctgagcggtatt 4840
Qy 3311 ttctccttacgcatctgtgcggtatttcacgcgcataatggcgcatctcagtcacaatct 3370
Db 4841 ttctccttacgcatctgtgcggtatttcacgcgcataatggcgcatctcagtcacaatct 4900
Qy 3371 gctgtatgcgcgcatagtttaagcagcccgacaccccgcaacaccccgctgacgcgcct 3430
Db 4901 gctgtatgcgcgcatagtttaagcagcccgacaccccgcaacaccccgctgacgcgcct 4960
Qy 3431 gacgggcttgcgtcctccggcatccggttacagacaagctgtgacccgtctccgggagct 3490
Db 4961 gacgggcttgcgtcctccggcatccggttacagacaagctgtgacccgtctccgggagct 5020
Qy 3491 gcatgtgcagaggttttcacgcgtacacccgaaacgcgcgagacgaagggcctcgtga 3550
Db 5021 gcatgtgcagaggttttcacgcgtacacccgaaacgcgcgagacgaagggcctcgtga 5080
Qy 3551 tacgcctattttatagggttaatgtcatgataataatggtttcttagacgtcaggtggca 3610
Db 5081 tacgcctattttatagggttaatgtcatgataataatggtttcttagacgtcaggtggca 5140
Qy 3611 cttttcgggaaatgtgcgggaacccctatttatttttttataaaacacattcaataa 3670
Db 5141 cttttcgggaaatgtgcgggaacccctatttatttttttataaaacacattcaataa 5200
Qy 3671 tgtatccgctcatgagacaataaacctgataaatgcttcaataataattgaaaaaggaaga 3730
Db 5201 tgtatccgctcatgagacaataaacctgataaatgcttcaataataattgaaaaaggaaga 5260
Qy 3731 gtatgagtattcaacatttccggtgccttattcccttttttttggcgattttgcttc 3790
Db 5261 gtatgagtattcaacatttccggtgccttattcccttttttttggcgattttgcttc 5320
Qy 3791 ctgttttctcaccacccagcaacgcgtggtgaaagtaaaagctgtaagatcgtgaggtgg 3850

QY 4571 cctcaactgattgaagcattgtaactgtcagaacaaagtcttactcatatatactatttagattg 4630
DB 6605 CCTCACTGATTAAAGCATTTGTAATCTGTGACACCAAGTTTACTCATATATATCTTTAGATTG 6546
QY 4631 atttaaaactctctttttaaattgaaaggtactagtggaagatcccttttttgataactctca 4690
DB 5545 ATTTAAANCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATATCTCA 6486
QY 4691 tgacaaaatcccttaacgtgagttcttcctcaactgagcgtcagaccccgtagaaaaaga 4750
DB 6485 TGACCAAAATCCCTTAACGTAAGTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGA 6426
QY 4751 tcaaaagatctcttgagatccctttttttctgcggttaactctgctctgcaacaaaaa 4810
DB 6425 TCAAGGATCTTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTCAACAAAAA 6366
QY 4811 aacacgcctcaccagcggtgtgtttgttccggtacaaagtaccaaactcttttccga 4870
DB 6365 AACACCGCTACCAAGCGGTGTTGTTGCGGATCAAGAGCTACCAACHTCTTTTCCGA 6306
QY 4871 agttaactggcttcagcagcgcagataccaaaatactgtccttctagttagtcgtagt 4930
DB 6305 AGCTAACTGGCTTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGT 6246
QY 4931 taggcacacactcaagaactctgtagcaccgcctacataactcgtctcgtactaactcgt 4990
DB 6245 TAGGCCACCACTTCAAGAACTGTGTAGCACCGCTTACATACCTCGCTCTGCTAATCCTGT 6186
QY 4991 taccagtggtcgtgcagtgccgataagtcgtcttaccgggttgacctcaagcagat 5050
DB 6185 TACCACTGGCTGCTGCAGTGGCGGATAAGTCTGTCTTACCGGGTTGGACTCAAGAGAT 6126
QY 5051 agttaccgataaaggcgagcgtcgggtgaaacgggggttcgtgcacacagccccagct 5110
DB 6125 AGTTACCGGATAAGGCGCAGCGTCGGCTGAACGGGGGTTCTGTCACACAGCCACGCT 6066
QY 5111 tgagagcaacacactcacacgaactagatacctacagcgtgagctatgagaaagcgca 5170
DB 6065 TGGAGCAAGGAGCTACCGAACTCAGATPACCTACAGCGTGAGCTATGAGAAAGGCCCA 6006
QY 5171 cgtctccgaaggagaaagcgccaggtatccggttaagcgcagggctcggaacagagag 5230
DB 6005 CGCTTCCGNAAGGAGANAGCGGACAGGTATCCGGTAAGCGCAGGGTCGGAACAGGAG 5946
QY 5231 agcgcagagggtctccagggggaacgcctggtatctttatagtcctgtcggtttc 5290
DB 5945 AGCGCACGAGGAGCTTCCAGGGGGAACGCCCTGATCTTTATAGTCTCTGCGGGTTC 5886
QY 5291 gccacctctgactgagcgtcgatttttgtatgctcgtcagggggcgagcctatgga 5350
DB 5885 GCCACCTCTGACTTGACGGTGCATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA 5826
QY 5351 aaaaacccagcaacgcgcttttttcggttctcggtccttttctggccttttgcctcaaca 5410
DB 5825 AAAACGCCAGCAGCGGCTTTTACGGTTCTTGCCCTTTTCTGGCCTTTTGTCTCACA 5766
QY 5411 tgtcttctcgtggtatcccttgatctctgttgataaccggtattaccgctttgagtgag 5470
DB 5765 TGTCTCTCTCGCTTATCCCTGATCTCTGTGATAACCGTATTACCGCCTTTTGTAGTGAG 5706
QY 5471 ctgataccgtctccagcagccgaacacagcagcagcagcagcagcagcagcagcagcag 5530
DB 5705 CTGATACCGCTCGCCGACGCGCAACACGAGCGCAGCGAGTCAGTCAGCGAGGAACGG 5646
QY 5531 aagagcgcccaatacgcgaacacccctctcccgcgcttgccgattcattaatgcagct 5590
DB 5645 AAGAGCGCCCAATACGCAAAACGCCCTCTCCCGCGGCTTGGCCGATTCATTATGCAGCT 5586
QY 5591 ggcaacacaggtttcccgactggaaagcgggcgagtgagcgcaacgcaaatatgtgagtt 5650
DB 5585 GGCACACAGGTTTCCCGACTGGAAAGCGGCGCAGTAGGCGCAACGCAATTAATGTGAGTT 5526
QY 5651 agctcaactcaataggcacccccaggctttacacactttatgcttcgggctcgtatgttg 5710

DB 5525 AGCTCACTCATTTAGGCCACCCAGGCTTTACACTTTATGCTTCGGGCTCGTATGTTGTG 5466
QY 5711 gaattggagcgcgatacaatttcacacagaaacagctatgaccatgattacgcca 5767
DB 5465 GAATTTGAGCGGATACAATTTTCACACAGAAACAGCTATGACCATGATTACGCCA 5409
RESULT 8
AAD04945/c
ID AAD04945 standard; DNA; 8153 BP.
XX AAD04945;
XX DT 17-JUL-2001 (first entry)
XX Plasmid pRK74 used to test Cre recombinase mediated inversion.
XX DE Gene trapping construct; conditional mutation; unidirectional inversion;
KW recombinase recognition sequence; RRS; disruption cassette;
KW selection cassette; transgenic organism; pRK74; Cre recombinase; ds.
XX OS Chimeric - Adenovirus.
OS Chimeric - Unidentified.
OS Chimeric - ECMV virus.
XX PN W0200129208-A1.
XX PD 26-APR-2001.
XX PF 16-OCT-2000; 2000MO-EP10162.
XX PR 16-OCT-1999; 99EP-0120592.
XX PR 27-OCT-1999; 99US-0162016.
XX PA (ARTE-) ARTEMIS PHARM GMBH.
XX PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX PI Kuehn R, Von Melchner H, Altschmied J;
XX WPI; 2001-308486/32.
XX PT New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX Example 3; Page 68-70; 78pp; English.
XX CC The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is pRK74 vector, which is used to test Cre
CC recombinase mediated inversion, which is related to the invention.
XX Sequence 8153 BP; 1976 A; 2133 C; 2067 G; 1977 T; 0 other;
SQ

Query Match 45.7%; Score 2637; DB 22; Length 8153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3131 cgaattcaactgcccgtcgctttttacaaactgctgactgggaaaaacctgctgcttaccacaac 3190

Db 8136 CGAATCTACTGGCGCTGCTTTTACAACTCGTGACTGGGAAACCTGGCGTTACCCAAC 8077
Qy 3191 ttaatcgcttgagcacatcccccttcgcagcgtggcgttaataagcgaagggccgcga 3250
Db 8076 TTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAAATAGCAGAGGCCGCGCA 8017
Qy 3251 ccgatcgcccttcccaacagttgcgagcctgaaatggcgaatggcgccgtgatgcggtatt 3310
Db 8016 CCGATCGCCCTTCCCAACAGTTGCGACAGCTGAATGGCGAATGGCGCCTGATGCGGTATT 7957
Qy 3311 ttctccttacacatctgtgcggtatttcacaccgcacatgatgactcctcagtaaatct 3370
Db 7956 TTCCTCTTACGCATCTGTGCCGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCT 7897
Qy 3371 gctctgatgcgcgatagtttaagcagcggcccccgcaccccgcccaacaccccgctgacgcgccc 3430
Db 7896 GCTCTGATGCGGCATAGTTAAGCCAGCGCCGACACCCGCCAACCCGCTGAGGGGCCCT 7837
Qy 3431 gacgggcttgctgctcccgccatccgcttacagacaagctgtgacctctccgggagct 3490
Db 7836 GACGGCTTGTCTGCTCCCGGCATCCCGTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 7777
Qy 3491 gcatgtgcagaggttttcacgctcatcacccgaacgcgcgagacgaagggccctcgtga 3550
Db 7776 GCATGTGTCAGAGTTTTCACCGTCATCACCGAAGCGCGGAGAGAAAGGGCCTCGTGA 7717
Qy 3551 tacgocctattttataggttaagtcaatgataataatggtttctcttagcgtcaggtggca 3610
Db 7716 TACGCCTATTTTATAGTTAATGTATCATGATAAATAAGTTCTTCTTAGACGTCAGGTGCA 7657
Qy 3611 cttttcggggaaatgtgcggcggaacccctatttggttatttttcttaataatacatcaata 3670
Db 7656 CTTTTCGGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTTAAATACATTTCAATA 7597
Qy 3671 tgatccgctcatgagacaataaacctgataaaatgcttcaataataatgaaaaaggaaga 3730
Db 7596 TGTATCCGCTCATGAGACAAATACCCGTATTAATGCTTCAATATATTTGAAAAGGAAGA 7537
Qy 3731 gtaagtgattcaacatttcgctgctgccccttattcccttttttggcgacattttgcccctc 3790
Db 7536 GTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTCGGGCATTTTGCCTTC 7477
Qy 3791 cgtgttttgcctaccagaaacgctggtgaaagttaaaagatgctgaagaatcagttgggtg 3850
Db 7476 CTGTTTTTGCCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 7417
Qy 3851 cacgagtggtttacatcgaaactggatctcaacagcggtaagatccttgagagtttcgcc 3910
Db 7416 CACGAGTGGGTTACATCGAACTGGATCTCAACAGGGGTAAAGTCTTAGAGATTTCGCC 7357
Qy 3911 ccgaagaacgttttccaatgatagacacttttaagttctgctgctggtggtggtattat 3970
Db 7356 CCGAAGAAGCTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGTATTAT 7297
Qy 3971 cccattatgcgcgggaagacaacactcgctcgccgacatacactattctcagaatgact 4030
Db 7296 CCGTATTGACGCGCGGCAAGAGCAACTCGGTCGCGGCATACACTATTCTCAGAAATGACT 7237
Qy 4031 tgggttgactcaccagtcacagaaaagcatcttacggaatggcgatgacagtaagagaat 4090
Db 7236 TGGTTGAGTACTACCAGTCAACAGAAAGCATCTTACGGATGGCATGACATGAGAGAAAT 7177
Qy 4091 tatcgagtctgcccataaccatgagtgaatacactgpgcccaacttactctgacaacga 4150
Db 7176 TATGAGTGTCTGCCATAACCATGAGTGATAACACTGCGGCCCAACTTACTCTGACAACGA 7117
Qy 4151 tcggaagaccgaagagtaaccgcttttttgacaaacatgggggagcatgataactcgc 4210
Db 7116 TCGGAGGACCGAAGAGGATTAACCGCTTTTTCACAACATGGGGGATCATGTAACCTGCC 7057
Qy 4211 ttgatcgttgggaacccgagctgaatgaagccatacacaacagcagcgtgacaccagga 4270

Db 7056 TTGATCGTGTGGNAACGGAGCTGAATGAAGCCCATACCAAAACAGAGCGTGACACACGA 6997
Qy 4271 tgcctgtagcaatggcaacaacgcttgcgcaaaacttaactgaggaacttaacttactctag 4330
Db 6996 TGCCGTGTAGCAATGCGCAACAAGTTGCGCAAACTATTAACTGCGCAACTACTTACTCTAG 6937
Qy 4331 ctteccggcaacaaataatagactggaatggaggcggaataaagttgcaggaacactctgc 4390
Db 6936 CTTCCGGCAACAATAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACTTCTGC 6877
Qy 4391 gctcggcccttccgctgctggtttatttctgataaaactctgagccggtgagcgtgggt 4450
Db 6876 GCTCGGCCCTTCCGCTGCTGCTGTTATTGCTGTATAAATCTGGAGCCGCTGAGCGTGGGT 6817
Qy 4451 ctgcggttatcatctgacgactggggccagatggtaagccctcccgtatcgtagttatct 4510
Db 6816 CTCGGGTATCATTTGACACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6757
Qy 4511 acacagcgggagtcaggcaactatggatgaacgaaaaatagacagatcgcctgagatagg 4570
Db 6756 ACACAGCGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGTAGGTG 6697
Qy 4571 cctcactgatgaagcattggttaactgtcagaccaggtttactcatatatacttttagctg 4630
Db 6696 CCTCACTGATTAAGCATTTGGTAACCTGTAGACCAAGTTTACTCATATATATCTTTAGATTG 6637
Qy 4631 atttaaaacttcatttttaatttaaaagatctagtggaagatcctcttttgaatactca 4690
Db 6636 ATTTTAAACTTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATATCTCA 6577
Qy 4691 tgaccaaaatcccttaacgctgaagtttctgctccactgagcgtcagaccccgtagaaaaa 4750
Db 6576 TGACAAAATCCCTTAAAGCTGAGTTTTCGTTCCACTGAGCGGTAGACCCCGTAGAAAAGA 6517
Qy 4751 tcaagagctctcttgagatccttttctgcgctgaatctgctgctgcaacaaaaa 4810
Db 6516 TCAAAAGGATCTCTTGAGATCCTTTTTTCTCGCGGTAACTGCTGCTTGCAAAACAAAA 6457
Qy 4811 aaccccgctaccagcgtggtttgttgcgggatcaagagctaccaactctttttccga 4870
Db 6456 AACCACCGCTACACAGCGGTGTTGTTTGC CGGATCAAGAGCTACCAACTCTTTTCCGA 6397
Qy 4871 aggttaacttgcttcagcagcgcagataacaaaaactgctcctctctagtgtagccgtagt 4930
Db 6396 AGGTAACTGGTTCAGCAGCGCAGATACCAAACTACTGCTTCTTAGTGTAGCCGTAGT 6337
Qy 4931 taggcacacactcaagaactctgtagccgcctacatacctcgtcgtctgctaaactct 4990
Db 6336 TAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTTAATCCTGT 6277
Qy 4991 taccagtgctgctccagtgggcgataagctgcttaccgggttggaactcaaacagat 5050
Db 6276 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGCTTTACCGGTTGGACTCAAGACGAT 6217
Qy 5051 agttaccggataaagcgcagcgttcgggctgaacgggggggttcgtgcacacagccagct 5110
Db 6216 AGTTACCAGTAAGCGCAGCGGTGCGGTGAACCGGGGGTTCTGTCACACAGCCAGCT 6157
Qy 5111 tggagcgaacgacctcacccgaactgagatacctcagcgtgagctatgagaagcgca 5170
Db 6156 TGGAGCAACGACCTTACCCGNACTGAGATACCTACAGCGTGAGTATGAGAAAGCGCCA 6097
Qy 5171 cgtctccgaagggaagcggagcaggtatccggtaagcggcggttcggaacagggag 5230
Db 6096 CGCTTCCGGAAGGGAGAAAGGGGACAGGTATCCGCTAAGCGCAGGGTCGGAACAGGAG 6037
Qy 5231 agcgcagaggaggaggttccagggggaacgcctggtatctttatagtcctctgcgggttcc 5290
Db 6036 AGCGACAGAGGAGGTTCCAGGGGAAACGCTGCTGATCTTTTATAGTCTGTCGGGTTTC 5977
Qy 5291 gscactctgacttgagcgtgatttttctgtagtctgtagggggggcgagcgttatgga 5350
Db 5976 GCCACTCTGACTTGAAGGCTCGATTTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGA 5917

QY 3671 tgtatcgcgtcatagagacaataaaccttgataaaatgcttcaataatatgaaaaaggaga 3730
Db 3680 tgtatcgcgtcatagagacaataaaccttgataaaatgcttcaataatatgaaaaaggaga 3639
QY 3731 gtatgagtattcaacattccgtgctgcgccttattcccttttttgcggcattttgccttc 3790
Db 3640 gtatgagtattcaacattccgtgctgcgccttattcccttttttgcggcattttgccttc 3699
QY 3791 ctgtttttgctacccagaaaacgtggtgaaagttaaagatgctgaagatcagttgggtg 3850
Db 3700 ctgtttttgctacccagaaaacgtggtgaaagttaaagatgctgaagatcagttgggtg 3759
QY 3851 cacagatgggttacatcgaaactggtatccaacagcggtaagatccttgagagttttcgcc 3910
Db 3760 cacagatgggttacatcgaaactggtatccaacagcggtaagatccttgagagttttcgcc 3819
QY 3911 ccgaagaacgttttccaatgatgagcacttttaaagtctctgtatgtggcgggtatttat 3970
Db 3820 ccgaagaacgttttccaatgatgagcacttttaaagtctctgtatgtggcgggtatttat 3879
QY 3971 cccgtattgacgcgggcaagagcaactcgtgctgccgcatacactatttctcagaatgact 4030
Db 3880 cccgtattgacgcgggcaagagcaactcgtgctgccgcatacactatttctcagaatgact 3939
QY 4031 tgggttgactactaccagtcagaaaaagacattcttaaggatggcatgacagtaagagaat 4090
Db 3940 tgggttgactactaccagtcagaaaaagacattcttaaggatggcatgacagtaagagaat 3999
QY 4091 tatgcagtgcctgcacataaccatgagtataacactgcggccaaacttacttctgacaacga 4150
Db 4000 tatgcagtgcctgcacataaccatgagtataacactgcggccaaacttacttctgacaacga 4059
QY 4151 tcggaggaccgaaggagtaaccgctttttgcacaacatggtgggagatcatgtaactcgcc 4210
Db 4060 tcggaggaccgaaggagtaaccgctttttgcacaacatggtgggagatcatgtaactcgcc 4119
QY 4211 ttgatcgttggaaaccgagtgatgaatgaagcattacaaacagacagcgtgacacacaga 4270
Db 4120 ttgatcgttggaaaccgagtgatgaatgaagcattacaaacagacagcgtgacacacaga 4179
QY 4271 tgcctgtagcaatggcaacaactgtgcgaactatttaacttgcgggaacttacttactctag 4330
Db 4180 tgcctgtagcaatggcaacaactgtgcgaactatttaacttgcgggaacttacttactctag 4239
QY 4331 ctcccggaacaataatagactggtatgagggcggaataaagtgcaggaacacttctgc 4390
Db 4240 ctcccggaacaataatagactggtatgagggcggaataaagtgcaggaacacttctgc 4299
QY 4391 gctcggcccttcggctggtggtttattgctgataaaatctggagccggtgagcgtgggt 4450
Db 4300 gctcggcccttcggctggtggtttattgctgataaaatctggagccggtgagcgtgggt 4359
QY 4451 ctgcggttatcatgtcagcactggggccagatggttaagccctcccgatatcgtagttact 4510
Db 4360 ctgcggttatcatgtcagcactggggccagatggttaagccctcccgatatcgtagttact 4419
QY 4511 acagacgggggagtcaggcaactatggtgaacgaataagacagatcgtcagatagtg 4570
Db 4420 acagacgggggagtcaggcaactatggtgaacgaataagacagatcgtcagatagtg 4479
QY 4571 cctcaactgattaaagcatlgttaactgtcagaccaagtttactcatatactattagattg 4630
Db 4480 cctcaactgattaaagcatlgttaactgtcagaccaagtttactcatatactattagattg 4539
QY 4631 atttaaaacttcaatttttaatttaaaggatctaggtgaagatcctttttgataatctca 4690
Db 4540 atttaaaacttcaatttttaatttaaaggatctaggtgaagatcctttttgataatctca 4599
QY 4691 tgacaaaatcccttaacgtgagtttcttccactgagcgtcagaccccgtagaaaaaga 4750
Db 4600 tgacaaaatcccttaacgtgagtttcttccactgagcgtcagaccccgtagaaaaaga 4659

QY 4751 tcaaaagattcttcttgagatccttttttctgcggttaactctgctgcttgcacacaaaaa 4810
Db 4660 tcaaaagattcttcttgagatccttttttctgcggttaactctgctgcttgcacacaaaaa 4719
QY 4811 aaccacgcctaccagcgtggtttgtttgcgggatacaagagctaccaactctttttccga 4870
Db 4720 aaccacgcctaccagcgtggtttgtttgcgggatacaagagctaccaactctttttccga 4779
QY 4871 aggttaactgggttccagcagcgcagataaccaaaaactctcttcttagtagccgtagt 4930
Db 4780 aggttaactgggttccagcagcgcagataaccaaaaactctcttcttagtagccgtagt 4839
QY 4931 taggcccacacttcaagaactctgtagcacgcctacatacctcgtctgtctaaacctgt 4990
Db 4840 taggcccacacttcaagaactctgtagcacgcctacatacctcgtctgtctaaacctgt 4899
QY 4991 taccagtggctgctccagtgccgataagtctgtcttaccgggttggactcaaacagat 5050
Db 4900 taccagtggctgctccagtgccgataagtctgtcttaccgggttggactcaaacagat 4959
QY 5051 agttaccgggataaaggcgcagcgtcgggtgaaacgggggttcgtgcacacagccagct 5110
Db 4960 agttaccgggataaaggcgcagcgtcgggtgaaacgggggttcgtgcacacagccagct 5019
QY 5111 tggagcgaacactacacacgaactgagatacctacagcgtgagctatgagaaagcga 5170
Db 5020 tggagcgaacactacacacgaactgagatacctacagcgtgagctatgagaaagcga 5079
QY 5171 cgcttcccgaagggaagacggcggaacaggtatccggttaagcggcagggctcggaacagag 5230
Db 5080 cgcttcccgaagggaagacggcggaacaggtatccggttaagcggcagggctcggaacagag 5139
QY 5231 agcgacagaggagcttccagggggaaacgctggtatctttatagtcctgtcgggtttc 5290
Db 5140 agcgacagaggagcttccagggggaaacgctggtatctttatagtcctgtcgggtttc 5199
QY 5291 gccactctgactgagcgtcgattttgtgatgctgcagggggggagccctatgga 5350
Db 5200 gccactctgactgagcgtcgattttgtgatgctgcagggggggagccctatgga 5259
QY 5351 aaaaacgcagcaacggccttttaccggttccgcttttgcggccttttgcctttgcacaa 5410
Db 5260 aaaaacgcagcaacggccttttaccggttccgcttttgcggccttttgcctttgcacaa 5319
QY 5411 tgttttctcgtgcttatccctgattctgtggaataacogtattaccgcctttgagtgag 5470
Db 5320 tgttttctcgtgcttatccctgattctgtggaataacogtattaccgcctttgagtgag 5379
QY 5471 ctgataccgctcgcgcagcgcagcaacgcagcgcagcgcagcgcagcgcagcgcagcgcg 5530
Db 5380 ctgataccgctcgcgcagcgcagcaacgcagcgcagcgcagcgcagcgcagcgcagcgcg 5439
QY 5531 aagagcgcccaatacgcacaaacgcctctcccgcgcttcccgcttcccgcttcccgct 5590
Db 5440 aagagcgcccaatacgcacaaacgcctctcccgcgcttcccgcttcccgcttcccgct 5499
QY 5591 ggcacgcaggtttcccgactggaagcggcagtgagcgcagcgcagcgcagcgcagcgcagcgcag 5650
Db 5500 ggcacgcaggtttcccgactggaagcggcagtgagcgcagcgcagcgcagcgcagcgcagcgcag 5559
QY 5651 agctcaactattagcaccacggctttacacttttatcttctcgcgctgattgtgtg 5710
Db 5560 agctcaactattagcaccacggctttacacttttatcttctcgcgctgattgtgtg 5619
QY 5711 gaattgtgagcgcgaatacaaatcttcacaggaacagcgtatgacattgattcgcca 5767
Db 5620 gaattgtgagcgcgaatacaaatcttcacaggaacagcgtatgacattgattcgcca 5676

RESULT 10
AAV63734
ID AAV63734 standard; DNA; 9335 BP.
XX

FT	CDS	5865..6419	
FT		/tag= s	
FT		/product= phosphinothricin acetyltransferase gene	
FT		/note= "Basta resistance selectable marker"	
FT	3'UTR	6420..6699	
FT		/tag= t	
FT		/note= "nos 3'UTR"	
FT	misc_feature	6700..9335	
FT		/tag= u	
FT		/note= "pUC19 sequences"	
PN	WO9856921-A1.		
XX	17-DEC-1998.		
XX	10-JUN-1998;	98WO-US11921.	
XX	12-JUN-1997;	97US-0049752.	
XX	(DOWC) DOW AGROSCIENCES LLC.		
XX	Ainley M, Armstrong K, Belmar S, Folkerts O, Hopkins N;		
XX	Menke MA, Pareddy D, Petolino JF, Smith K, Woosley A;		
XX	WPI; 1999-080904/07.		
XX	New isolated regulatory sequences for transgenic plants - which are		
XX	derived from the maize root preferential cationic peroxidase protein		
XX	(per5) gene		
XX	Example 17; Page 117-122; 150pp: English.		
XX	This is the nucleotide sequence of PGP/367, a plasmid containing		
CC	the maize root preferential cationic peroxidase per5 gene (see		
CC	(see AA6317) promoter, the per5 untranslated leader modified to		
CC	include the per5 intron 1, the GUS gene, and the per5 3'		
CC	untranslated region (3'UTR). Because intron flanking sequences		
CC	(exon DNA) are important in the processing of the intron, 16 bases		
CC	of flanking exon DNA were included in the fusion within the per5		
CC	untranslated leader. Transformation experiments in maize		
CC	demonstrated that the presence of the per5 intron is essential		
CC	for root-specific expression from the per5 promoter. The invention		
CC	relates to isolated regulatory sequences, especially promoter,		
CC	intron and 3'UTR sequences, of the maize per5 gene. Claimed		
CC	recombinant gene cassettes comprising per5 regulatory sequences are		
CC	used to control expression of recombinant genes in selected tissue,		
CC	especially the root, of transformed plants, particularly maize.		
XX	Sequence 9335 BP; 2402 A; 2290 C; 2258 G; 2385 T; 0 other;		
XX	Query Match	45.7%; Score 2633.2; DB 20; Length 9335;	
XX	Best Local Similarity	99.9%; Pred. No. 0;	
XX	Matches 2635; Conservative	0; Mismatches 3; Indels 0; Gaps	
Qy	3130	gcgaattactgcccgcgtgtttatacaacgtctgactgggaaaacctggcggtaccacaa	3189
Db	6698	gggaattactgcccgcgtgtttatacaacgtctgactgggaaaacctggcggtaccacaa	6757
Qy	3190	cttaatgcgcttcgacacatcccccttcgcagctggcgtaatagcgaagagccgcgc	3249
Db	6758	cttaatgcgcttcgacacatcccccttcgcagctggcgtaatagcgaagagccgcgc	6817
Qy	3250	accgatgcgcccctcccaacagttgcgcagcctgaatggcgagctgatcggtat	3309
Db	6818	accgatgcgcccctcccaacagttgcgcagcctgaatggcgagctgatcggtat	6877
Qy	3310	ttctccttacgcatctgtcggtatttcacaccgcatatggtgcactctcagtcacaac	3369
Db	6878	ttctccttacgcatctgtcggtatttcacaccgcatatggtgcactctcagtcacaac	6937
Qy	3370	tgctctgatgcgcgatagtttaagcagagcccgacaccgcgcaaacaccgctgacgcgcc	3429


```
QY 5590 tggcagcagaggtttcccgactggaagcgggcagtgagcgaacgaataatgtgagt 5649
Db 9158 tggcagcagaggtttcccgactggaagcgggcagtgagcgaacgaataatgtgagt 9217
QY 5650 tagctcaactcattagcaccgccggtttacactttatgcttccggctcgtatgtgtgt 5709
Db 9218 tggctcaactcattagcaccgccggtttacactttatgcttccggctcgtatgtgtgt 9277
QY 5710 ggaattgtgagcggaatacaaatccacaggaagaaacagcagctatgacctatgacgcga 5767
Db 9278 ggaattgtgagcggaatacaaatccacaggaagaaacagcagctatgacctatgacgcga 9335

RESULT 11
AAV63742
ID AAV63742 standard; DNA; 6898 BP.
XX
AC AAV63742;
XX
DT 12-APR-1999 (first entry)
XX
DE Plasmid UGP232-4 containing ubi promoter and per5 3'UTR.
XX
KW Peroxidase; per5 gene; maize; corn; transgenic plant; promoter;
KW root; vector; plasmid UGP232-4; ubiquitin; ds; circular; cyclic.
XX
OS Chimeric - Escherichia coli.
OS Chimeric - Zea mays.
OS Chimeric - maize streak virus.
OS Chimeric - Agrobacterium tumefaciens.
OS Chimeric - synthetic.
XX

FH Key Location/Qualifiers
FT misc_feature 1..14
FT /tag= a
FT /note= "pUC19 polylinker"
FT promoter 15..993
FT /tag= b
FT /note= "maize ubiquitin promoter"
FT intron 994..2007
FT /tag= c
FT /note= "ubiquitin intron"
FT misc_feature 2008..2026
FT /tag= d
FT /note= "synthetic linker"
FT CDS 2027..3835
FT /tag= e
FT /product= beta-glucuronidase
FT misc_feature 3836..3890
FT /tag= n
FT /note= "sequence from pKA882"
FT 3'UTR 3897..4262
FT /tag= o
FT /note= "per5 3'UTR nt 6066-6431"
FT misc_feature 4269..6898
FT /tag= p
FT /note= "pUC19 sequence"
XX
XX W09856921-A1.
XX
XX 17-DEC-1998.
XX
XX 10-JUN-1998; 98WO-US11921.
XX
XX 12-JUN-1997; 97US-0049752.
XX
XX (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Ainley M, Armstrong K, Belmar S, Folkerts O, Hopkins N;
XX Menke MA, Pareddy D, Petolino JF, Smith K, Woosley A;
XX WPI; 1999-080904/07.
XX
```

```
PT New isolated regulatory sequences for transgenic plants - which are
PT derived from the maize root preferential cationic peroxidase protein
XX (per5) gene
XX
XX Example 21; Page 126-129; 150pp; English.
XX
CC This is the nucleotide sequence of UGP232-4, a plasmid comprising
CC a gene cassette in which the following components are operably
CC joined: the maize ubiquitin (ubi) promoter, the GUS gene and the
CC maize root preferential cationic peroxidase per5 gene (see AAV63717)
CC 3' untranslated region (3'UTR). Testing in maize callus indicated
CC that the per5 3'UTR functioned 19% better than nos when the GUS
CC gene was driven by the ubi promoter. This could result from
CC changes in the efficiency of processing or increased stability of
CC the message. The invention relates to isolated regulatory
CC sequences, especially promoter, intron and 3'UTR sequences, of the
CC maize per5 gene. Claimed recombinant gene cassettes comprising
CC per5 regulatory sequences are used to control expression of
CC recombinant genes in selected tissue, especially the root, of
CC transformed plants, particularly maize.
XX
SQ Sequence 6898 BP; 1705 A; 1617 C; 1680 G; 1896 T; 0 other;
```

```
Query Match 45.7%; Score 2632.8; DB 20; Length 6898;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2634; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3132 gaattcactggcgcgtgttttacaacgtcgtgactgggaaacccctggcgtaccacaact 3191
Db 4263 gaattcactggcgcgtgttttacaacgtcgtgactgggaaacccctggcgtaccacaact 4322
QY 3192 taatgccttgcagcacatccccctttccagctggcgttaatagcgaagagccgcac 3251
Db 4323 taatgccttgcagcacatccccctttccagctggcgttaatagcgaagagccgcac 4382
QY 3252 cgatgccttcccaacagttgcgcagcgtgaatggcgaatggcgcgtgattt 3311
Db 4383 cgatgccttcccaacagttgcgcagcgtgaatggcgaatggcgcgtgattt 4442
QY 3312 tctcctcagcatctgctggtatttcacacgcgcacatggtgcactctcagtacacatcg 3371
Db 4443 tctcctcagcatctgctggtatttcacacgcgcacatggtgcactctcagtacacatcg 4502
QY 3372 ctctgatgcgcgcatagtttaagccagcccgacaccccgccacaccccgctgacgcgcctg 3431
Db 4503 ctctgatgcgcgcatagtttaagccagcccgacaccccgccacaccccgctgacgcgcctg 4562
QY 3432 acgggcttctgctcccgccgcttccagcagaagctgtgacgcgtctccgggagctg 3491
Db 4563 acgggcttctgctcccgccgcttccagcagaagctgtgacgcgtctccgggagctg 4622
QY 3492 catgtgtcagaggttttcacccgtcatcacgaaacgcgcgagacgaaggcctcgtgat 3551
Db 4623 catgtgtcagaggttttcacccgtcatcacgaaacgcgcgagacgaaggcctcgtgat 4682
QY 3552 acgcctatttttataggttaatgtcatgataataatgggtttcttagacgtcaggtgggcac 3611
Db 4683 acgcctatttttataggttaatgtcatgataataatgggtttcttagacgtcaggtgggcac 4742
QY 3612 ttttcgggaaatgtgcgggaacccctatttattttatttttcaataacattcaaatat 3671
Db 4743 ttttcgggaaatgtgcgggaacccctatttattttatttttcaataacattcaaatat 4802
QY 3672 gtatccgctcatgagacaataacccctgataaatgcttcaataatattgaaaaggagag 3731
Db 4803 gtatccgctcatgagacaataacccctgataaatgcttcaataatattgaaaaggagag 4862
QY 3732 tatgagtattcaacatttccgctgctccctatttcccttttttgcggcattttgccttcc 3791
Db 4863 tatgagtattcaacatttccgctgctccctatttcccttttttgcggcattttgccttcc 4922
QY 3792 tgtttttgtcaccacgaacacgctggtagaagtaaaagatgctgaagatcagttgggtgc 3851
```


Qy	4633	ttaaaacttcaatttttaatttaaaggatctaggtgaagatacctttttgataatctcatg	4699
Db	2610	ttaaaacttcaatttttaatttaaaggatctaggtgaagatacctttttgataatctcatg	2669
Qy	4693	accaaaatcccttaacgtgagtttcgttccactgagcgtcagaccocgtagaaaagatc	4752
Db	2670	accaaaatcccttaacgtgagtttcgttccactgagcgtcagaccocgtagaaaagatc	2729
Qy	4753	aaaggatctcttgagatcctttttctgcggttaactctgctctgcaaacacaaaaa	4812
Db	2730	aaaggatctcttgagatcctttttctgcggttaactctgctctgcaaacacaaaaa	2789
Qy	4813	ccaccgctaccagcgggtgttctgttcgggatcaagagctaccaaactcttttccgaag	4872
Db	2790	ccaccgctaccagcgggtgttctgttcgggatcaagagctaccaaactcttttccgaag	2849
Qy	4873	gtaacctggcttcagcagagcgagataccaaatactgctctctctagctagccgtagta	4932
Db	2850	gtaacctggcttcagcagagcgagataccaaatactgctctctctagctagccgtagta	2909
Qy	4933	ggccaccacttcaagaactctgacacgcgtacataacctcgctctgctaaactctgta	4992
Db	2910	ggccaccacttcaagaactctgacacgcgtacataacctcgctctgctaaactctgta	2969
Qy	4993	ccagttggctgctgcagtgcgcataagtcgtgtcttaccgggttggactcaagacgatag	5052
Db	2970	ccagttggctgctgcagtgcgcataagtcgtgtcttaccgggttggactcaagacgatag	3029
Qy	5053	ttaccggataaaggcgacgcggtctgggctgaaacgggggttcgtgcacacagccacgttg	5112
Db	3030	ttaccggataaaggcgacgcggtctgggctgaaacgggggttcgtgcacacagccacgttg	3089
Qy	5113	gagcgacgacctacacgaactgagataccttaccagcgtgagctatgagaaaagccacg	5172
Db	3090	gagcgacgacctacacgaactgagataccttaccagcgtgagctatgagaaaagccacg	3149
Qy	5173	cttcccgaaaggagaaaggcgacaggtatccggttaagcgaggggtcgaaacagggag	5232
Db	3150	cttcccgaaaggagaaaggcgacaggtatccggttaagcgaggggtcgaaacagggag	3209
Qy	5233	cgacagagggagcttccaggggaaacgcttggtatctttatagctctcggttttcgc	5292
Db	3210	cgacagagggagcttccaggggaaacgcttggtatctttatagctctcggttttcgc	3269
Qy	5293	caactctgacttgagctcgatttttggagctcgtcagggggcgagagcctatggaaa	5352
Db	3270	caactctgacttgagctcgatttttggagctcgtcagggggcgagagcctatggaaa	3329
Qy	5353	aacgcagcaaacggcccttttaaggttcttgcccttttgcgttgcccttttgcacatg	5412
Db	3330	aacgcagcaaacggcccttttaaggttcttgcccttttgcgttgcccttttgcacatg	3389
Qy	5413	ttctttccctcggttatccctgatctgtgataaccgtattaccgccttttgatgaact	5472
Db	3390	ttctttccctcggttatccctgatctgtgataaccgtattaccgccttttgatgaact	3449
Qy	5473	gataccgctcgcgcagccgaacacgcagcgacgcagtcagtgagcgagggaagcgaaa	5532
Db	3450	gataccgctcgcgcagccgaacacgcagcgacgcagtcagtgagcgagggaagcgaaa	3509
Qy	5533	gagcgcccaatacgcaaacgcctctcccccgcggttgcccgattcaatgcagctgg	5592
Db	3510	gagcgcccaatacgcaaacgcctctcccccgcggttgcccgattcaatgcagctgg	3569
Qy	5593	cacgacaggtttcccgacttgaaaacggcgagtgagcgcaacgcaattaatgtgagttag	5652
Db	3570	cacgacaggtttcccgacttgaaaacggcgagtgagcgcaacgcaattaatgtgagttag	3629
Qy	5653	ctcaactataggcacccacgctttacactttatgctctccgctcgatgtgtgtgga	5712
Db	3630	ctcaactataggcacccacgctttacactttatgctctccgctcgatgtgtgtgga	3689
Qy	5713	attgtgagcggatacaacaaatttcacacaggaacagctatgacctgattacgcaca	5767

Db	3690	attgtgagcgataacaattttcacacaggaacacgctatgaccatgattacgccca	3744
RESULT	13		
ABA04130			
ID	ABA04130	standard; DNA; 3755 BP.	
XX			
AC	ABA04130;		
XX			
DT	28-FEB-2002	(first entry)	
XX			
DE	Plasmid pHCEI9(II)	polynucleotide sequence SEQ ID NO:4.	
XX			
KW	Promoter; gene expression; fermentation; D-AAT; Bacillus SK-1;		
KW	D-amino acid aminotransferase; economic; large-scale protein production;		
KW	circular; ds.		
XX			
OS	Bacillus sp. SK-1.		
OS	Synthetic.		
XX			
EH	Key	Location/Qualifiers	
FT	Promoter	1..225	
FT		/*tag= a	
FT		/note= "specifically claimed promoter region in	
FT		Claim 1 (see ABA04128)"	
XX			
PN	WO200183787-A1.		
XX			
PD	08-NOV-2001.		
XX			
PF	26-APR-2001; 2001WO-JP03607.		
XX			
PR	27-APR-2000; 2000JP-0128528.		
XX			
PA	(BIOL-) BIOLEADERS CORP.		
PA	(TAKI-) TAKARA SHUZO CO LTD.		
XX			
PI	Sung M, Lee S, Hong S, Seo H;		
XX			
DR	WPI; 2002-066535/09.		
XX			
PT	Promoter sequence effective in Escherichia coli and Bacillus for		
PT	economic large-scale fermentative production of proteins -		
XX			
PS	Claim 7; Page 58-60; 65pp; Japanese.		
XX			
CC	The present invention describes a DNA sequence (I) comprising promoter		
CC	activity in Escherichia coli or Bacillus cells, where the promoter is		
CC	derived from the D-AAT (D-amino acid aminotransferase) gene of		
CC	Bacillus SK-1. (I) can be used for the high level expression of a		
CC	foreign gene in a bacterial host for economic and efficient large-scale		
CC	production of proteins such as enzymes, cytokines and antibodies. The		
CC	present sequence represents the plasmid pHCEI9(II) polynucleotide		
CC	sequence comprising a specifically claimed promoter region from the		
CC	present invention.		
XX			
SQ	Sequence 3755 BP; 951 A; 914 C; 932 G; 958 T; 0 other;		
Query Match 45.6%; Score 2631.8; DB 24; Length 3755;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps			
Qy	3133	aattcaactggccgtcggtttttacaacgtcgtagctgggaaacccctggcgttacccaactt	3192
Db	1112	aattcaactggccgtcggtttttacaacgtcgtagctgggaaacccctggcgttacccaactt	1171
Qy	3193	aatcgcttgcagacatccccctttgcgcagctggcgtaataagcgaagggccgcacc	3252
Db	1172	aatcgcttgcagacatccccctttgcgcagctggcgtaataagcgaagggccgcacc	1231
Oy	3253	qatcgcccttcccaacgatttcgcagcgtgaatagcaatggccgcctgacgtatattt	3312

Qy	5473	gatacgcctcgccgagcccgaacgaccgagcgcgagtcagtgagcggaaacgcgcaa	5533
Df	3452	gatccgctctcgccgagccccgaacgaccgagcgcgagtcagtgagcggaaacgcgcaa	3511
Qy	5533	gagcgcccccaatcacgaaaacgcctctccccgcgcgttggccgattcattaatgcagctgg	5592
Df	3512	gagcgcccccaatcacgaaaacgcctctccccgcgcgttggccgattcattaatgcagctgg	3571
Qy	5593	cacgacaggtttcccgactggaaagcggcagtgcgacgcaacgcgaattaatgtgagttag	5652
Df	3572	cacgacaggtttcccgactggaaagcggcagtgcgacgcaacgcgaattaatgtgagttag	3631
Qy	5653	ctcaactatgagcaccgcccgactttacaccttatgtcttcgcgctcgtatgtgtgtgga	5712
Df	3632	ctcaactatgagcaccgcccgactttacaccttatgtcttcgcgctcgtatgtgtgtgga	3691
Qy	5713	attgtgagcgggataacaatttcacacagagaacacagctatgaccatgattacgcca	5767
Df	3692	attgtgagcgggataacaatttcacacagagaacacagctatgaccatgattacgcca	3746
RESULT 14			
ID	AAT86449/C		
AC	AAT86449 standard; DNA; 4283 BP.		
CC	AAT86449;		
CT	10-FEB-1998 (first entry)		
XX	DNA encoding hGH, EGF fusion protein from pWRG1630.		
KW	Mature epidermal growth factor; EGF; wound healing; skin; gold bead;		
KW	epidermal cell; hair follicle; polyA addition sequence;		
KW	human growth hormone secretory signal peptide; hGH; ds.		
OS	Chimeric - homo sapiens.		
OS	Chimeric - bos taurus.		
OS	Synthetic.		
Key	Location/Qualifiers		
CDS	713..1253		
FT	/tag= a		
FT	/note= "contains intronic sequences"		
sig_peptide	713...1049		
FT	/tag= b		
FT	/standard_name= secretion_signal		
FT	/note= "human growth hormone secretory signal contains intronic sequences"		
mat_peptide	1050..1250		
FT	/tag= c		
FT	/product= mature_human_EGF		
exon	713...723		
FT	/tag= d		
Intron	/number= 1		
FT	724..982		
FT	/tag= e		
FT	/number= 1		
FT	983..1253		
FT	/tag= f		
FT	/number= 2		
FT	/note= "end of exon is given as TAG stop codon"		
XX	US5661132-A.		
PX	26-AUG-1997.		
XX	22-NOV-1994;	94US-0343401.	
XX	14-DEC-1989;	89US-0451957.	
PR	22-MAY-1991;	91US-0707248.	
PR	11-JUN-1992;	92US-0897357.	
PR	11-JUN-1993;	93US-0076550.	

|||||
Db 3665 ATGAGTATTCAACATTTCCGCTGCGCCCTTATTCCCTTTTTCGGCATTTTGCCTTCCT 3066
QY 3793 gtttttgcctaccagaaacgcgtgggtgaaagttaaagatcgtctgaagatcagttgggtgca 3852
Db 3605 GTTTTGGCTCACCAGAAACGCTGGTGAAGTGAAGATGCTGAAGATCAGTTGGGTGCA 3546
QY 3853 cgagtggtttacatcgaactggaatcgaacagcgggtaagatcctttgagagttttcgcccc 3912
Db 3545 CGAGTGGGTTACATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 3486
QY 3913 gaagaacgttttccaatgatgacacacttttaaaagtctctgatactggtcgcggtattatcc 3972
Db 3485 GAAGAAGGTTTCCATATGATGACACTTTTAAAGTTTCTGCTATGTGGCGGGTATTATCC 3426
QY 3973 cgtattgacccggggaagagcaactcgggtcgccgatacaactattcttcgaatgacttg 4032
Db 3425 CGTATTGACCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGATGACTTG 3366
QY 4033 gttgagtactcaccagtcacagaaaaagcatcttcacggatggcgatgaagtaagaatta 4092
Db 3365 GTTTCAGTACTCAGCAGTTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA 3306
QY 4093 tgcagtgctccataaccatgagtataacactgcgccaacttaacttcttgacaagatc 4152
Db 3305 TGCAGTGTCTGCCATACCATGATGATGAACACTGCGGCCCAACTTACTTCTGACAACGATC 3246
QY 4153 ggaggaccgaagagctaaacgcgttttttgcacaaatgggggatcatgttaactcgcctt 4212
Db 3245 GGAGGACCGAAGGAGCTAACCGCTTTTTCGCACACATGCGGGATCATGTAACTCGCCTT 3186
QY 4213 gatcgttgggaacccgagctgaatgaagccataccaaacacagcagcgtgacacaagatg 4272
Db 3185 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACAGATG 3126
QY 4273 cctgtagcaatgcacaacgcttgcgaactatttaactggcgaactacttacttagct 4332
Db 3125 CTTGTAGCAATGCACACGTTTGGCAAACTATTAACTGGCGAATCTTACTCTAGCT 3066
QY 4333 tcccggaacaataatagactggatggagcggaataaagtgcaggaaccacttctgcgc 4392
Db 3065 TCCCGGCAACAATAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACTTCTGCGC 3006
QY 4393 tcggcccttcggctggctgtttattgctgataaaatcgtggagccgggtgagcgtgggtct 4452
Db 3005 TCGGCCCTTCCCGCTGGCTGGTTTATGCTGAATAATCTGGAGCCGGTGAGCGTGGGTCT 2946
QY 4453 cgcggtatcatcagcactggggccagatggtgaagccctcccgatcgtagttatctac 4512
Db 2945 CGCGGTATCATTCAGCACTTGGGGCCAGATGGTAAGCCCTCCCGTATCTGTAGTTATCTAC 2886
QY 4513 acgacggggagtcaggcaactatggatgaacgaaatagacagatcgtctgagatagtgcc 4572
Db 2885 ACGACGGGAGTCAGCAACTATGGATGAACGAATAAGACAGATCCTGAGATAGGTGCC 2826
QY 4573 tcaatgataagcatcggtaactcgtcagaccaagtttactcatatatactttagatgat 4632
Db 2825 TCACATTAAGCATTTGGTAACGTGTGAGAACCAAGTTTACTCATATATACCTTTAGATTGAT 2766
QY 4633 ttaaaacttcatcttaatttaaaggatcctaggtgaagatccttttgaataatctcatg 4692
Db 2765 TTAATACTTCATTTTAAATTAAGAGATCTAGGTGAAGATCCCTTTTGTAAATCTCATG 2706
QY 4693 accaaatacccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaagatc 4752
Db 2705 ACCAAAATCCCTTAACGTGAGTTTTCGTCCACTGAGCGTGCAGACCCCGTAGAAAAGATC 2646
QY 4753 aaaggatcttcttgagatcctttttctcgcggtaatctgctgcttgcaaaaaa 4812
Db 2645 AAAGGATCTCTCTGAGATCCTTTTTCGCGGCTAATCTGCTGCTTGCAACAAAAA 2586
QY 4813 ccaccgctacacgcggtgggtttgttgcggatcaagagctaccaactcttttccgaag 4872
|||||

Db 2585 CCACCGCTACCACGGGTGGTTTGTTCGCCGATCAAGAGCTACCAACTCTTTTCCGAAG 2526
QY 4873 gtaactggcttcagcagcagcgcagataccaaataactgtctcttagtgtagccgtagttta 4932
Db 2525 GTAACGGCTTCAGCAGAGCGCAGATACCAAAATACGTCTCTTCTAGTGTAGCCGTAGTTA 2466
QY 4933 ggcacacacttcaagaactcgttagcaccgcctacataacactcgtctctgctaaatcctgta 4992
Db 2465 GGCACACACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTTAATCTGTTA 2406
QY 4993 ccaatggctctccagtggcgataagtctgtctaccgggttggaactcaacagatag 5052
Db 2405 CCAGTGGCTCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTGGACTCAAGACCATAG 2346
QY 5053 ttaccggataagcgcagcggctcggtgaacgggggttcgtgcacacagccacagcttg 5112
Db 2345 TTACCGGATAAGCGCAGCGTGGCTGAACGGGGGTTCTGTGCACACACCCAGCTTG 2286
QY 5113 ggcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcacg 5172
Db 2285 GAGCGACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCACG 2226
QY 5173 cttcccgaaaggagaagcgcagcaggtatccggtgaacggcgaggttcggaacagagag 5232
Db 2225 CTTCCCGAAGGGAGAAGCGGACAGGTATCCGGTAAGCGCGCAGGGTTCGGAACAGGAGAG 2166
QY 5233 cgcacgagggagcttccagggggaacgcctggtatctttatagtcctctcggttttcgc 5292
Db 2165 CGCACGAGGGAGCTTCCAGGGGAAACGCCCTGTATCTTTATAGTCTCTGTCGGGTTTCGC 2106
QY 5293 caactctgacttgagcgtcgatattttgtgtagctcgtcgaagggggcgagccatgaaaa 5352
Db 2105 CACTCTGACTTGAGGCTCGATTTTGTGATGCTGCTCAGGGGGGCGAGCCATATCGAA 2046
QY 5353 aacgcagaaacgcgctttttacggttctgctgcttttgcgcttttgccttttgcctcatg 5412
Db 2045 AACGCCAGCAACCGCGCTTTTACGCTTCTGGCTTTTGTCTGGCTTTTGTCTCATG 1986
QY 5413 ttcttctcgttatccctgattctgtgataacgcgtattaccgccttttgagtgagct 5472
Db 1985 TTCTTTCTCGCTTATCCCTGATTTCTGTGATAACCGTATTACCCTTTTGTAGTGAGCT 1926
QY 5473 gatacgcctcgcgcagcgcgaacagcagcagcagcagcagcagcagcagcagcagcagc 5532
Db 1925 GATACCGCTCGCGCAGCGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1866
QY 5533 gagcgcccaatagcgaacacgcctctcccgcgcttcccgcgcttccatgaatgcagctgg 5592
Db 1865 GAGCGCCCAATAGCAAAACCGCCTCTCCCGCGCGTGGCCGATTCATTAAATGCAGCTGG 1806
QY 5593 cagcagaggtttcccgactggaaagcggcgagcgcgaacgaatgaatgagttag 5652
Db 1805 CACGACAGGTTTCCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTGAGTTAG 1746
QY 5653 ctaactcattagcagcccccaggcttttaactttatgcttccggtcgtatggttggga 5712
Db 1745 CTCACATCATTAGCACCCCGAGGCTTTACACTTTATGCTTCCGCTCGTATGTTGTGGA 1686
QY 5713 attgtgagcggataaacaatttcacaggaacagcatgacccatgattacgca 5767
Db 1685 ATTGTGAGCGGATACAAATTTACACAGGAAACAGCTATGACCTATGATACGCCA 1631

RESULT 15
AAT86610/C
ID AAT86610 standard; DNA; 4283 BP.
XX
AC AAT86610;
XX
DT 26-MAR-1998 (first entry)
XX
DE Epidermal growth factor-encoding plasmid DNA pWRG1630.
XX

KW Epidermal growth factor; injection; microneedle; delivery;
XX genetic material; skin cell; wound; protein expression; ds.
OS Synthetic.
OS Homo sapiens.

XX Key Location/Qualifiers
FH 713..1253
FT CDS /*tag= a
FT /note= "contains intron"
FT 723..981
FT Intron /*tag= b
XX

XX US5697901-A.

XX 16-DEC-1997.

XX 19-MAY-1995; 95US-0445265.

XX 14-DEC-1989; 89US-0451957.

XX 22-MAY-1991; 91US-0707248.

XX 11-JUN-1992; 92US-0897357.

XX 11-JUN-1993; 93US-0076550.

XX (ERIK/) ERIKSSON E.

XX Eriksson E;

XX WPI; 1998-051325/05.

XX P-PSDB; AA443092.

XX Introducing genetic material into tissue cells - by injection with
XX microneedle

XX Example 2; Columns 15-20; 21pp; English.

XX The present sequence represents epidermal growth factor-encoding plasmid
CC DNA which was used in a new method for introducing genetic material
CC into a cell, comprising repeatedly injecting the genetic material into
CC cells of a target tissue at an in-situ site with a microneedle, whereby a
CC change in the animal attributable to the delivery of the genetic material
CC is detectable. The target tissue is selected from skin cells of a wound,
CC and the target cells may be covered with a protective chamber. The
CC method is used for obtaining long-term expression of native or non-native
CC proteins, e.g. for secretion into the bloodstream or lymphatic system, or
CC for preventing expression of proteins eliciting an immune response.

XX SQ Sequence 4283 BP; 1023 A; 1095 C; 1114 G; 1051 T; 0 other;

Query Match 45.6%; Score 2631.8; DB 19; Length 4283;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3133 aattcaactggcgtcttttaacacgtctgactggaaacccctgggttacccaactt 3192
DB 4265 AATTCACCTGGCGCTGCTTTTACACGTCTGACTGGGAACCCCTGGCGCTTACCCAACCT 4206
OY 3193 aatgccttgcagcacatccccctttccagctggcgttaataagcgaagagcccgacc 3252
DB 4205 AATCGCCTTGAGACACATCCCCCTTTTCGCACGTGGCGTAATAGCGAAGAGCCCGCACC 4146
OY 3253 gatcgcccttcccaacagtctgcgagcctgaatggcgaatggcgcctgatgcggtatttt 3312
DB 4145 GATTCGCCCTTCCCAACAGTTGGCAGCCTGAATGGCGAATGGCGCTGATGCGGTAATTTT 4086
OY 3313 ctccctacgcactctgctggtattttcacaccgcacatatggtcactctcagtcacacatctgc 3372
DB 4085 CTCCTTAGCATCTGTGGGGTATTTTCACACCGCATATGGTGACACTCTAGTACAACTCTGC 4026
OY 3373 tctgatgcgcacatagtttaagccagcccgacaccccgcaacaccccgctgacgcgcctga 3432
DB 4025 TCTGATCGCGCATAGTTAAGCCAGCCCGACACACCGCCCAACACCCGCTGACCGCGCCTGA 3966

OY 3433 cggggttctctgctcccggaataccgcttacagacaagctgtgacogtctccgggagctgc 3492
DB 3965 CGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGAGCTGC 3906
OY 3493 atgtgtcagaggttttcaccgtcatcacccgaacacgctgagacgaagcctcgata 3552
DB 3905 ATGTGTCAGAGGTTTTTCACCGTCAATCACCAGAACCGCCGAGAGAAAGGCCCTCGTGATA 3846
OY 3553 cgcctattttataggttaattgctcatgataataaagtttcttagcgtcagtggtgcaact 3612
DB 3845 CGCCTATTTTATAGTTAATGTCATGATAAATAGTGTTCCTTAGACGTGAGTGGCACT 3786
OY 3613 ttccggggaaatgtgcgggaacccctaattgtttatttttcttaataacattcaataatg 3672
DB 3785 TTTCCGGGAAATGTGCGGAAACCCCTATTGTGTTTATTATTTTCTAATAATACATTCAAATATG 3726
OY 3673 tatccgtcatgagacaataaacctgataaattgcttcaataattgataaataaaggaagat 3732
DB 3725 TATCCGCTCATGAGACATAAACCTGATAAATGCTTCAATAATATTTGAAAGGAAGAGT 3666
OY 3733 atgagtattcaacatttccgtgtcgcccttattcccttttttggcgcattttgccttcc 3792
DB 3665 ATGAGTATTCAACATTTTCGCTGTCGCCCTTATTCCCTTTTTCGGGCATTTTTCCTTCC 3606
OY 3793 gttttgtcccccgaacacgctgtggaagtaaaagatgcgtgaagatcagttggtgca 3852
DB 3605 GTTTTGTCTACCCAGAAACGCTGTTGTAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 3546
OY 3853 cgaatgggttacatcgaaactgattctcaacagcgtgaagctcttgagagtttccccc 3912
DB 3545 CAGTGGGTTACATCAAACTGGATCTCAACAGCGTAAAGATCTTGGAGAGTTTTCGCCCC 3486
OY 3913 gaagaacgttttccaatgatgagcacttttaagtttctctatgtggtggttatctac 3972
DB 3485 GAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCGCGGTATTATCC 3426
OY 3973 cgtattgacgcgggcaagacaactcggctgcgcgcatacacatactattctcagaatgacttg 4032
DB 3425 CGTATTGACCGCGGCAAGACCACTCGGTGCGCGCATACACTATTCTCAGATGACTTGT 3366
OY 4033 gttgagtactccacgtcacaaaaagcatcttcagatggtgcatgacatagagaatta 4092
DB 3365 GTTGAGTACTACACAGTACAGAAAGCATCTTTACGGATGGCATGACATGAAGAGAAATTA 3306
OY 4093 tgcagtgtgcataaccatgagtgataacactgcggccaacttacttctgcacaacgac 4152
DB 3305 TGCAGTGTCTGCCATAACCATGAGTGATAACACTGCGGCCCACTTACTTCTGACAAGATC 3246
OY 4153 ggagacacgaagagctaacccgtttttgcacaaactcgtggggatcgttaactcgcctt 4212
DB 3245 GGAGGACCGAAGAGCTAACCCGCTTTTTCACAACTATTTTTCACATGGGGATCATGTAACTCGCCT 3186
OY 4213 gatccttgggaacccgagctgaatgaagcataaccacgaacgagcgtacacacagatg 4272
DB 3185 GATCGTTGGGAACCGAGCTGAATGAAGCCATACCAACAGACGAGCGTGACACACAGATG 3126
OY 4273 cctgtagcaatggcaacaacgttgcgcaactatttaactggcggaactacttactctagct 4332
DB 3125 CCTGTAGCAATGCCAACACGTTGCGCAAACTATTAACTGGCGAAGTACTTACTCTAGCT 3066
OY 4333 tccccggcaacaattaatagactggatggaggcggaataaagtgtcaggaacactcttcgac 4392
DB 3065 TCCCGCAACAATTAATAGACTGGATGGAGGGGATTAAGAGTTTGCAGAGCACTTCTGCGC 3006
OY 4393 tgggcccctccggctggctggtttattgctgataaatctggagccggtgagcgtgggtct 4452
DB 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCCGGTGAGCGTGGTCT 2946
OY 4453 cgcggtatcttgacgactgggcccagatggttaagccctcccgtatcgttagttatctac 4512
DB 2945 CCGGCTATCATTCAGCACTGGGGCCAGATGTTAACCCCTCCCGTATCTGAGTTATCTAC 2886

QY 4513 acagcggggagtcagcgaactatggatgaacgaataatagacagatcgctgagatgagtgcc 4572
|||||
Db 2885 ACAGCGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGTGCC 2826
QY 4573 tcaactaaagcatggtaactgcagacccaagtttaactatatactatttagattgat 4632
|||||
Db 2825 TCACTGATTAAAGCATGGTAACCTGTGAGACCAAGTTTACTCATATATACCTTTAGATTGAT 2766
QY 4633 ttaaaacttcatttttaatttaaaagatctaggtaagaatcccttttttgataatctcatg 4692
|||||
Db 2765 TTAATACTTCATTTTAAATTTAANAAGATCTAGGTGAAGATCCCTTTTGTATATCTCATG 2706
QY 4693 accaaatcccttaacgtgagtttctccactgagcgtcagaccccgtagaaaagatc 4752
|||||
Db 2705 ACCAAAATCCCTTAAGCTGAGTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATC 2646
QY 4753 aaagatctcttgagatccctttttctgcgcgtaactgtcgttgcaacaaaaaa 4812
|||||
Db 2645 AAAGGATCTTCTTGAGATCCCTTTTTCGCGGTAACTGTGCTTGCATTCGCAACAAAAA 2586
QY 4813 ccacgcctacacaggggtgtgttgcgcggtacagagctacaaactctttttccgaag 4872
|||||
Db 2585 CCACCGCTACACAGGCTGGTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAG 2526
QY 4873 gtaactggcttcagcagagcgagataccaaaatactgtcctcttagtgtagccgtagtta 4932
Db 2525 GTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCGGTAGTTA 2466
QY 4933 ggcacacactcaagaactctgtagcacgcctacatacctcgtctgttaactcctgta 4992
|||||
Db 2465 GGCACACACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGTCTGTAAATCCTGTTA 2406
QY 4993 ccagtgtcgtccagtgcgagatgaagtctgtcttaccgggttgactcaagacgatatag 5052
Db 2405 CCAGTGGCTGCTCCAGTGGCGATAGTCGTGTCTTACCGGGTGGACTCAAGACGATAG 2346
QY 5053 ttaccggataagcgcagcgttcggcgtgaacgggggttcgtgcacacagccagcttg 5112
Db 2345 TTACCGGATAAAGCGCAGCGGTGGGCTGAACGGGGGTTCGTGCACACAGCCAGCTTG 2286
QY 5113 gagcgaacgacctacacgaactgagatacctacagcgtgagctatgagaaagcgcacg 5172
Db 2285 GACGGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCACG 2226
QY 5173 ctcccgaaagggaagcgagacaggtatccggtaagcggcaggggtcggaaacagagag 5232
Db 2225 CTTCCGGAAGGGAAGAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAG 2166
QY 5233 cgacgagggagcttcacgggggaacgcctgttatcttatagtcctgtcgggtttcgc 5292
Db 2165 CGCACGAGGAGAGCTTCAGGGGGAACGCCTGTGTATCTTTATAGTCTCTGCGGGTTTCGC 2106
QY 5293 cactctgacttgagcgtcgatttttgatgctcgtcagggggcgagcctatgaaa 5352
Db 2105 CACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGCGGACCCATATCGAAA 2046
QY 5353 aacgcagcaacgcggcctttttacgggttcctggccttttctggccttttctgctcacatg 5412
Db 2045 AAGCCAGCAACCGCGCTTTTACGGTTCCTGGCTTTTGTGCTGGCTTTTGTCTACATG 1986
QY 5413 ttcttctcgtgtatccccctgattctgtgataaacctgattaccgcctttgagtgagct 5472
Db 1985 TTCTTCTCTCGGTATCCCTGATTCTGTGGATTAACCGTATTACCGCCTTTGAGTGAGCT 1926
QY 5473 gatccgctcgccgcagccgaacgcagcagcagcagcagcagcagcagcagcagcagcag 5532
Db 1925 GATACCGCTCGCCGACCGCAACGACCGGAGTCAGTGACCGAGGAGGAGCGGAA 1866
QY 5533 gagcgcacaaatagcgaacccgtctccccgcgcttggccgatttcattaatgacgtgg 5592
Db 1865 GAGCGCCCAATAGCAAAACCGCTCTCCCGCGCTTGGCGGATTTCATTAATGCAGCTGG 1806
QY 5593 cagcagaggtttcccgactggaaagcggcgagtcgagcgcaacgcaattaatgtgagttag 5652

Db 1805 CACGACAGGTTTCCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAG 1746
QY 5653 ctcaactattagcaccacccaggttttacacttttatgcttccgactcgtatgttgga 5712
Db 1745 CTCACTATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTGTGGA 1686
QY 5713 attgtgagcggataaacaatttcacacaggaacagcgtatgacctgattacgcca 5767
Db 1685 ATTGTGAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 1631

Search completed: August 31, 2002, 18:25:58
Job time: 16960 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:24:39 ; Search time 6064.26 Seconds
(without alignments)
12835.357 Million cell updates/sec

Title: US-09-810-861b-3
Perfect score: 5767
Sequence: 1 agcttgatgctgcaggtc.....ctatgaccatgattacgcga 5767

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_estl: *
10: gb_est2: *
11: gb_htc: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vrt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1480.2	22.2	2080	11	BC001541
c 2	1055.8	18.3	1070	9	AJ281552
c 3	986.4	17.1	1013	10	BM438846
c 4	950.8	16.5	1067	9	AU081137
c 5	924	16.0	1004	9	AJ281480
c 6	919.4	15.9	1089	9	AU081124
c 7	875.8	15.2	935	10	BG38279
c 8	841	14.6	841	9	AL042026
c 9	824.6	14.3	854	10	BM438950
c 10	782.4	13.6	1163	9	AU081044
c 11	777.8	13.5	800	9	AJ281449
c 12	772.4	13.4	934	9	AL044364
c 13	764.2	13.3	990	10	BI489087
c 14	750.4	13.0	931	10	BI488594
c 15	743.6	12.9	832	10	BG923768
c 16	738.6	12.6	842	10	BI687610
c 17	722.4	12.5	724	9	AL645114

c 18	720.8	12.5	847	12	AZ687169
c 19	715.8	12.4	792	10	BI687057
c 20	711.8	12.3	715	9	AL661706
c 21	705.4	12.2	707	9	AL656688
c 22	704	12.2	818	10	BI180988
c 23	701.8	12.2	705	9	AL635845
c 24	701.8	12.2	705	9	AL635845
c 25	701.4	12.2	838	12	AL643164
c 26	698	12.1	718	10	AZ679307
c 27	697.8	12.1	780	10	AL631067
c 28	691	12.0	1039	9	BI753192
c 29	688.6	11.9	822	9	AU081040
c 30	686.4	11.9	689	9	AL190022
c 31	686	11.9	687	9	AL646532
c 32	680.6	11.8	700	9	AJ281616
c 33	679	11.8	680	9	AU001481
c 34	678	11.8	788	9	AL646751
c 35	677.8	11.8	757	12	AJ281697
c 36	673.2	11.7	754	10	4A3A-P9E1
c 37	672.4	11.7	693	9	AQ856202
c 38	671	11.6	693	9	BG920379
c 39	669	11.6	670	9	AL636713
c 40	668.4	11.6	670	9	AL598527
c 41	666.4	11.6	669	9	AL662130
c 42	665.8	11.5	672	9	AL635952
c 43	665.6	11.5	703	9	AL660789
c 44	664.8	11.5	671	9	AL640799
c 45	662.4	11.5	675	9	AJ281437
					AL660035
					AL645322

ALIGNMENTS

RESULT	1	BC001541	Homo sapiens, clone IMAGE:3453362, mRNA.	2080 bp	linear	HTC 31-JAN-2002
LOCUS	BC001541	Homo sapiens				
DEFINITION	BC001541	Homo sapiens				
ACCESSION	BC001541	Homo sapiens				
VERSION	BC001541.1	human				
KEYWORDS	HTC					
SOURCE	human					
ORGANISM	Homo sapiens					
REFERENCE	1	(bases 1 to 2080)				
AUTHORS	Strausberg, R.					
TITLE	Direct Submission					
JOURNAL	Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov					
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcdepaxil.stanford.edu R. McKon, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, D. M.					
FEATURES	Location/Qualifiers					
source	1. .2080 /organism="Homo sapiens"					

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 4 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156
This clone has the following problem: no polyA-tail.

QY 2468 tggacct 2474
 Db 2074 TGGACCT 2080

RESULT 2
 AJ281552/2

LOCUS 1070 bp mRNA linear EST 30-JUN-2000
 DEFINITION 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
 gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
 AJ281552
 VERSION 1
 KEYWORDS EST
 SOURCE African malaria mosquito.
 ORGANISM Anopheles gambiae

REFERENCE 1 (bases 1 to 1070)
 AUTHORS Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
 Donohue, M., Schultze, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.,
 and Kafatos, F.C.
 TITLE Anopheles gambiae pilot gene discovery project: identification of
 mosquito innate immunity genes from expressed sequence tags
 generated from immune-competent cell lines
 Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
 20300950
 COMMENT Contact: Dimopoulos G
 Fotis C. Kafatos laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany.
 FEATURES
 source Location/Qualifiers
 1..1070
 /organism="Anopheles gambiae"
 /strain="4A r/r"
 /db_xref="taxon:7165"
 /clone="4A3A-P6F11"
 /cclone.lib="Anopheles gambiae immune competent 4A3A"
 /cell_line="immune competent 4A3A"
 /lab_host="E. coli DH10B"
 /note="Vector: p7T3b-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
 forward priming site which reads from the 3' end of the
 cDNA. The 4A3A is a directionally cloned and normalized
 cDNA library that was constructed from the 4A3A cell line
 oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
 (1996) : Normalization and Subtraction: Two approaches To
 Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 263 a 283 c 255 g 269 t
 ORIGIN

Query Match 18.3%; Score 1055.8; DB 9; Length 1070;
 Best Local Similarity 99.7%; Pred. No. 9.8e-214;
 Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 4293 gttggcaaacattataactgaggaactacttacttagcttcccggaacaattataga 4352
 Db 1070 GTTGGCGCAAACTATTAACTGGCGAACTACTTACTTAGCTTCCCGGCAACAATTAATAGA 1011

QY 4353 ctggatggagcggaataagttgcagaccactctgcgctcgccctccggtcggtcg 4412
 Db 1010 CTGGATGAGCGGGATAAAGTTGACAGACCACCTTCTGGCGTCCG-CCTTCCGCGTGGCTG 952

QY 4413 gtttattgctataaatctggagcgggtgagcgtgggtctgcggtatcattgcagcact 4472
 Db 951 GTTTATTGCTGATAAATCTGGAGCGGGTGAGCGTGGGTCTCGCGTATCATTTGACGACT 892

QY 4473 gggggcagatggttaagccctccgctatcgtagttatctacagcagggagtcagggaac 4532
 Db 891 GGGGCGACATGTTAAGCCCTCCCGTATCGTAGTTATCTACAGAGCGGGAGTCAGGCAAC 832

QY 4533 tatgatgaacgaatagacagatcgctgagataggtccctcactgattaaagcattgta 4592

Db 831 TATGATGGAGAAATAGACAGATCGCTGAGATAGGTGCTCCTACCTGATTAAGCATTTGTA 772
 QY 4593 actgtcagaccagaagtttactcatatatacttttagattgatttaaaactctcattttaatt 4652
 Db 771 ACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTAAAT 712

QY 4653 taaaaggatctaggtgaagatcccttttgataatctatgcacaaaatcccttaacgtga 4712
 Db 711 TAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATGACCAAAATCCCTTTAAGCTGA 652

QY 4713 gttttcgttccactgagcgtcagaccgcgtagaaaaagatcaaaagatcttcttgagatcc 4772
 Db 651 GTTTTCGTTCCACTGAGCGTCAGACCCGCTAGAAAAGATCAAAAGATCTTCTTTGAGATCC 592

QY 4773 ttttttctgcgcgttaactctgctgtgtcgaacaaaaaacaccgcgtaccagcggtggt 4832
 Db 591 TTTTTCCTCGCGTAACTGCTGCTGCAACAAACAAAACCCGCTTACAGACGGTGGT 532

QY 4833 ttgtttgcgcatcaagagctaccacactcttttccgaaggtaactggtcctcagcagagc 4892
 Db 531 TTGTTTGGCGGATCAAGAGTACCAAACTCTTTTCCGAAGGTAAGTGGCTTCCAGCAGAGC 472

QY 4893 gcagatacacaatactgctccttcttagtagcgcgtagtaggceaccactcaagaactc 4952
 Db 471 GCAGATACCAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAATC 412

QY 4953 ttagcacccgcctacatacctcgtctctgctaatcctgtttaccagtggtcgtgccagtgg 5012
 Db 411 TGTAGCACCGCCTACATACCTCGCTCTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGG 352

QY 5013 cgataagtcgtctgtcttaccggttggactcaagacgtagttaccgggataaagcgacgg 5072
 Db 351 CGATAAGTCGTGCTTTACCGGGTTGGACTCAAGAGAGATAGTTACCGGATAAGCGCAGCG 292

QY 5073 gtccggctgaacgggggttcgtcacacagccagcttgagcgaacagacacacacga 5132
 Db 291 GTCGGGCTGACGGGGGGTTCGTGCACACAGCCCACTTGGAGCGAACACCTACACCGA 232

QY 5133 actgsagatacctacagcgtgagctatgagaaagcgcacgcgtctcccgaggaggaagagc 5192
 Db 231 ACTGAGATACCTACAGCGTGAGCTATGAGAAAGCCCGCTTCCGGAAGGGAAGGC 172

QY 5193 ggaacggtatccggttaacgagcgggttcggaacagagagcagcagggagctccagg 5252
 Db 171 GGACAGGATATCCGTTAAGCGGAGGGTGGAAACAGAGAGCGACAGGAGGAGCTTCCAGG 112

QY 5253 gggaacgcctcggtatctttatagtcctcgtcggtttccacacctcctgactgagcgtcg 5312
 Db 111 GGGAAACGCTGCTATCTTTATAGTCTCTGCGGGTTTCCGACCTCTGACTTGAGCGTGG 52

QY 5313 attttgtgctcgtcagggggcgagcctatgaaaaacgcagca 5363
 Db 51 ATTTTGTGATGCTCGTCAGGGGGCGGAGCGCTATGAAAAACGCCAGCAA 1

RESULT 3
 BM438846
 LOCUS 1013 bp mRNA linear EST 31-JAN-2002
 DEFINITION IPLVR00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
 sequence.
 BM438846
 VERSION 1
 KEYWORDS EST
 SOURCE channel catfish.
 ORGANISM Ictalurus punctatus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
 Ictaluridae; Ictalurus.
 1 (bases 1 to 1013)
 REFERENCE 1
 AUTHORS Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
 TITLE Transcriptome of channel catfish (Ictalurus punctatus): initial
 analysis of expressed sequence tags from the liver

JOURNAL
COMMENT

Unpublished (2002)

Contact: Liu ZJ

The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular BiosciencesAuburn University, Auburn, AL 36849, USA
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA

Tel: 334 844 4054

Fax: 334 844 9208

Email: zliu@acesag.auburn.edu

Seq primer: M13 Reverse.

Location/Qualifiers

1. .1013

/organism="Ictalurus punctatus"

/db_xref="taxon:7998"

/clone_lib="Liver cDNA library"

/note="Organ: Liver; Vector: pSport1; Site_1: NotI;

Site_2: Sali"

273 a 228 c 245 g 266 t 1 others

BASE COUNT
ORIGIN

Query Match 17.1%; Score 986.4; DB 10; Length 1013;

Best Local Similarity 99.5%; Pred. No. 5.5e-199;

Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 3778 gcaatttgcttcctctgttttgcaccagaaacgctggtgaaagtataaagatgctgaa 3837

Db 1 GCATTTTGCTTCCTGCTTTTGTCTACCCAGACGCTGGTGAAGTTAAAGATGCTGAA 60

QY 3838 gatcagttgggtcacagagtggtttacatcgaaactgagatcacaagcggtgaagatcctt 3897

Db 61 GATCAGTTGGGTGCAGAGTGGTTCATCGAACTGGATCTCAACAGCGGTAAAGATCCCT 120

QY 3898 gagagtttgcgcgcgaagaacgttttccaatgatgagcacttttaaaagtctgctatgt 3957

Db 121 GAGAGTTTGGCCCGGAGNAGACGTTTCCATGATGAGCACTTTTAAAGTTCGCTATGT 180

QY 3958 ggcgcggtattatcccgattgacgcgcgggaagagcaactcggcgccgacatacattat 4017

Db 181 GCGCGGTATTATCCGGTATTACGCGCGGCAAGAGCAACTCGTGGCGCATACACTAT 240

QY 4018 tctcagaatgactggttgagtagtactcacagtcacagaaagcatcttaccggtgcatg 4077

Db 241 TCTCAGAATGACTGTTGAGTGTACTCACCAGTCACAGAACGATCTTACGGATGGCATG 300

QY 4078 acagtaagagaattatgcagtgctgcataaaccatgagtatacaactgcgcgcaactta 4137

Db 301 ACAGTAAGAGAATTATGCAGTGTGCTGCCATAACCATGATGATAACACTGCGGCAACTTA 360

QY 4138 ctctcgacaacgactcgaggagaccgaaggagctaacgcgttttttcgacacaatcggggat 4197

Db 361 CTCTCACAACAGATCGGAGACCGCAA-GAGCTAACCGCTTTTGTGCAACATGGGGGAT 419

QY 4198 c-atgtaactcgccttgatcggttggaaacgagcgtgaatgaagccatcacaacagcga 4256

Db 420 CAATGTAACTCGGCTTGATCGTTGGGAACGGAGCTGAATGAAGCCATACCAAAACGACGA 479

QY 4257 gcgtgacaccagatgcctgtagcaatggcaacaacgctgtgcgaactattaaactggcga 4316

Db 480 GCGTGACACCACGATGCCGTAGCAATGCAACAACGTTGCGCAAACTATTAACTGGCGA 539

QY 4317 actactactctagtctcccgcaacaataatagactggatggagcggtataaagtgc 4376

Db 540 ACTACTTACTCTAGCTCCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGC 599

QY 4377 agaccacttctgcgtcgcgccttcgcgtggtgtttattgctgataaactctggagc 4436

Db 600 AGACCACATTTCTCGGCTCGGCCCTTCGGCTGGCTGGTTATTTCGTGATAAATCTGGAGC 659

QY 4437 cggtagcggtggtctcgcggtatcattgcagcactggggccagatggtaagccctcccg 4496

Db 660 CGGTGAGCGTGGGTCTCGCGGTATCATTCGAGCACTGGGGCCAGATGGTTAAGCCCTCCG 719

QY 4497 tatcgtagttatctacacagcggggagtcaggcaacatggtgtaacgaagaatagacagat 4556
|||||
Db 720 TATCGTAGTTATCTACACGACGCGGAGTCAGGCAACTATGATGAACGAAATAGACAGAT 779
|||||
QY 4557 cgtgagataggcgcctcactgattgaagcatgtgaactgtcagacaaagtattactcata 4616
|||||
Db 780 CGCTGAGATAGGTGCCTCACTGATTAAAGCATGTGTAACCTGTCAGACCAAGTTTACTCATA 839
|||||
QY 4617 tatacttagattgatttaaaacttcatttttaatttaaaagatctagtggaagatcct 4676
|||||
Db 840 TATACTTTAGATTGATTAAACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCT 899
|||||
QY 4677 ttttgataatctcatgacccaaatcccttaacgtgagtttcttcactgagcgtcaga 4736
|||||
Db 900 TTTTGTGATAATCTCATGACCAAAATCCCTTAACTGAGTTTTCGTTCCACTGAGCGTCAGA 959
|||||
QY 4737 ccccgtagaaagatcaaggatctctctgagatcctctttttctgcggttaat 4790
|||||
Db 960 CCCCGTAGAAAGATCAAGGRTCTTCTTGAGATCCTTTTTTCTGCGCGTAAT 1013
|||||

RESULT 4
AU081137/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

977; Conservative

0; Mismatches

7; Indels

2; Gaps

2;

QY 4782 ggcgcgtaactcgtctgttcgaaacaaaaaaccccgctaccagcggtgtgtgttgc 4841

Db 1067 GCGCGTAATCTGCTGC-TGCACAAACAAAAACACGCTACACAGC-GTTGTTTGTGGC 1010

QY 4842 ggaatcaagagctaccactcttttccgaaggttaactggttcagcagagcgagatacc 4901

Db 1009 GGATCAAGAGCTACCAACTCTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACC 950

QY 4902 aaatactgctctctctagtagtagcgcgttagtaggcccaccacttcaagaactcttagcacc 4961

Query Match

16.5%; Score 950.8; DB 9; Length 1067;

Best Local Similarity

99.1%; Pred. No. 2.1e-191;

Matches

977; Conservative

0; Mismatches

7; Indels

2; Gaps

2;

QY 4782 ggcgcgtaactcgtctgttcgaaacaaaaaaccccgctaccagcggtgtgtgttgc 4841

Db 1067 GCGCGTAATCTGCTGC-TGCACAAACAAAAACACGCTACACAGC-GTTGTTTGTGGC 1010

QY 4842 ggaatcaagagctaccactcttttccgaaggttaactggttcagcagagcgagatacc 4901

Db 1009 GGATCAAGAGCTACCAACTCTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACC 950

QY 4902 aaatactgctctctctagtagtagcgcgttagtaggcccaccacttcaagaactcttagcacc 4961


```
QY 4922 agcgtagtttagccaccacttcaagaactctgttagcaccgcctacataacctgctctgc 4981
|||||
Db 471 AGCGTAGTTAGGCCACCACCTCTGAAGAACTCTGTAGCACCGCCTACATACCTGCTCTGC 412
|||||
QY 4982 taatccctgttaccagtgctgcctgcagtgccgataagctgcttaccgggtggact 5041
|||||
Db 411 TAATCCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTGGACT 352
|||||
QY 5042 caagaagatagttaccggaataagcgcagcgtcgccggcgaacggggggttcgtgcacac 5101
|||||
Db 351 CAAGACGATAGTTACCGGATAAGCGCAGCGGCTCGGGCTGAACGGGGGTTCTGTGCANAC 292
|||||
QY 5102 agccagcttgagcgaacacactcacccgaactgagatacctacagcgtgagctatgag 5161
|||||
Db 291 AGCCAGCTTGGAGCGAACAACCTACACCGAAGTACCTACAGCGTGAGCTATGAG 232
|||||
QY 5162 aaagccacgcctcccgaaaggagaaagcgcagaggtatccggtaagcgcagggctg 5221
|||||
Db 231 AAAGCGCCACGCTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGCAGGGTCG 172
|||||
QY 5222 gaacagagagcgcagcagggagctctcaggggggaaacgcctggtatctttatagtcctg 5281
|||||
Db 171 GAACAGAGAGCGCACGAGGAGCTTCCAGGGGGAACGCTCGTATCTTTATAGTCCCTG 112
|||||
QY 5282 tcggggttcgccacctgtgactgagcgtgcgatttttgtgatgctcgtcagggggcgga 5341
|||||
Db 111 TCGGGTTTCGCCACCTCTGACTTTGAGCGTCGATTTTGTGATGCTGCTCAGGGGGCGGA 52
|||||
QY 5342 gctatgaaacacgcagcagcgcctttttacggcttcctggcctttt 5392
|||||
Db 51 GCCTATGGAANAACGCAGCAACGCGGCCCTTTTACGGTTCTTGGCCCTTTT 1
|||||

RESULT 6
AU081124/c 1089 bp mRNA linear EST 30-AUG-2001
LOCUS AU081124 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cdna clone KG'12,
mRNA sequence.
ACCESSION AU081124
VERSION AU081124.1 GI:6431472
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Kono,T., Sakai,M. and Lapetra,S.E.
TITLES Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakukenibana1, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.

FEATURES
source
1..1089
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KG'12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="kidney infected by infectious hematopoietic
necrosis virus"
245 a 297 c 295 g 251 t 1 others

BASE COUNT 245 a 297 c 295 g 251 t 1 others
ORIGIN
|||||
```

Query Match 15.9%; Score 919.4; DB 9; Length 1089;

RESULT 7

```
Best Local Similarity 97.3%; Pred. No. 9.9e-185;
Matches 978; Conservative 0; Mismatches 21; Indels 6; Gaps 4;
QY 4769 atccttttttctgcgcgttaactgc-tgcttgcacaacaaaaa---aacacacgcgtacc 4824
|||||
Db 1089 ATCCGTTTTTTTTCGCGGGTAATCTGCGTGTCTGCMAACAAAAGACACCCCTTCCAG 1030
|||||
QY 4825 gcgggtggttggttgocggat-caaagactaccacactcttttccgaagtaactggctt 4883
|||||
Db 1029 CGGTTGTTTTTGTGTCGGATCCAAGAGCTACCCACTCTTTTTCCGAAGTAACCTGGCTC 970
|||||
QY 4884 cagcagagcgcagataccaaaactgctccttctagtgtagccgttagtgccac-cact 4942
|||||
Db 969 CAGCAGAGCCAGATACAAAATATCTGTCTTCTAGTGTAGCCGTAGTAGCCACCAACT 910
|||||
QY 4943 tcaagaactctgtagcaccgcctacatacctcgtcctcgtctaactcctgttaccagtggtc 5002
|||||
Db 909 TCAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTPANTCCCTGTTACACAGTGGCTG 850
|||||
QY 5003 ctgccagtgcgataagtcgtgtcttaccgggttgactcaagacagatagttaccggata 5062
|||||
Db 849 CTGCCAGTGGCGATAAGTCTGTCTTACC GGTTGGACTCAAGACGATAGTTACCGGATA 790
|||||
QY 5063 agcgcagcgggtcggcgtgaacgggggttcgtgcacacagccacgcttggagcgaa 5122
|||||
Db 789 AGCGCAGCGGTTCGGCTGAACGGGGGGTTCGTGCACACAGCCACGCTTGGAGCGAACGA 730
|||||
QY 5123 cctacaccgaactgagatacctacagcgtgagctatgagaaagcgcacgttcccgaag 5182
|||||
Db 729 CTTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAG 670
|||||
QY 5183 ggagaagcgcagcaggtatccggtaagcggcgagggttcggaacagagagcgacaggg 5242
|||||
Db 669 GGAGAAAGCGCGACAGATATCCGGTAAGCGCAGGGTCGGAACAGGAGCGACGACGAGGG 610
|||||
QY 5243 agcttcagggggaaacgcctggtatctttagtctcgtcggggttccggcacctctgac 5302
|||||
Db 609 AGCTTCCAGGGGAAACGCTCGTATCTTTATAGTCTGTGCGGTTCGCGACCTCTGAC 550
|||||
QY 5303 ttgagcgtcgtattttgtgctcgtcagggggggcgagcctatgaaacacgcagca 5362
|||||
Db 549 TTGAGCGTCGATTTTGTGATGCTCGTACGGGGCGGAGCCCTATGMAAAGCCGACGA 490
|||||
QY 5363 acggcctttttacggctcctggtcctttgtcgccctttgtcacaatgttcttctg 5422
|||||
Db 489 AC CGCGCTTTTACGGCTTCTCGCCCTTTTGTGCTGCTTTCACATGCTTCTTCTG 430
|||||
QY 5423 cgttaccctgattctgtgataacgcgtattaccgcttaccgcttggagtgatagcgtc 5482
|||||
Db 429 CGTTATCCCTGATTCTGTGGATAACCGTATTACCGCTTTGAGTGAGCTGATACCGCTC 370
|||||
QY 5483 gccgcagcgaacgcagcgcagcgcagcgcagcgcagcgcagcgcagcgcagcgcagc 5542
|||||
Db 369 GC CGCAGCGCAACGACCGAGCGAGCGAGCTCAGTGAGCGAGGAGCGAGAGCGGCCAA 310
|||||
QY 5543 tacgaaacgcctctcccgccgcttggcggatttcattaatgagcgtggcagcaggt 5602
|||||
Db 309 TAGCAAAACCGCTCTCCCGCGCTTGGCGATTGCGGATTCATTAATGCAGCTGCGACGAGGT 250
|||||
QY 5603 ttcccgactggaagcggcgagcgcagcgaacgaataatgtgagttagctcaactt 5662
|||||
Db 249 TTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACCTATT 190
|||||
QY 5663 aggcacccaggttttacactttatgcttcggcgtcgtatgtgtgtggaattgtgagcg 5722
|||||
Db 189 AGCAGCCCGAGGCTTTACACTTTATGCTCCGCTCGTATGTTGTGGAAATGTGAGCG 130
|||||
QY 5723 gatacaatttcacaggaacacagctatgacctattacgcca 5767
|||||
Db 129 GATAACAATTTTCACAGGAAACAGCTATGACCATGATTACGCCA 85
|||||
```

```

BG838279          935 bp      mRNA      linear      EST 25-MAY-2001
LOCUS             Gc01_10e07_R Gc01_AAFc_ECORC_cold_stressed_Glycine_clandestina
DEFINITION        Glycine clandestina cDNA clone Gc01_10e07, mRNA sequence.
ACCESSION         BG838279
VERSION           BG838279.1 GI:14204601
KEYWORDS          EST
SOURCE            Glycine clandestina.
ORGANISM          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                  Glycine.
REFERENCE         1 (bases 1 to 935)
AUTHORS           Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A., Harris
                  ,L.J., Hattori,I.J., Ouellet,T., Robert,L.S., Sprott,D. and Tinker
                  ,N.A.
TITLE             Expressed Sequence Tags from Cold-Stressed Glycine clandestina
JOURNAL           Unpublished (2001)
COMMENT           Contact: Singh,J.A.
                  Eastern Cereal and Oilseed Research Centre
                  Agriculture and Agri-food Canada
                  960 Carling Avenue, Bldg. 20, Ottawa, Ontario, KIA 0C6, Canada
                  Tel: (613) 759-1662
                  Fax: (613) 759-1701
                  Email: singhja@em.agr.ca.
FEATURES          Location/Qualifiers.
                   1..935
                   /organism="Glycine clandestina"
                   /cultivar="1035"
                   /db_xref="taxon:45687"
                   /clone="Gc01_10e07"
                   /clone_lib="Gc01_AAFc_ECORC_cold_stressed_Glycine_clandest
                   ina"
                   /tissue_type="Leaves, stem"
                   /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
                   Site_2: XhoI; Plants incubated at 2 degrees under 12 hours
                   of light/day. Harvested after only 2-3 days of cold
                   treatment. cDNA was prepared with the Uni-Zap cDNA kit
                   from Stratagene. Eco RI adapters were linked followed by
                   digest with Xho I/Eco RI and ligated to pBluescript."
BASE COUNT        243 a 213 c 227 g 238 t 14 others
ORIGIN
...
Query Match      15.2%  Score 875.8;  DB 10;  Length 935;
Best Local Similarity 97.7%;  Pred. No. 1.8e-175;
Matches 882;  Conservative 14;  Mismatches 6;  Indels 1;  Gaps 1;

Qy  3593  cttagacgtcagggtggcactttctgggaaatgtgcgggaacccctattgtttatttt 3652
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  3653  tctaaatcattcaaatgtatcgcctcatgagacaataaccctgataaatgttcaat 3712
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  86    tctaataacattcaaatgtatcgcctcatgagacaataaccctgataaatgttcaat 145
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  3713  aatatgaaaaggagatgtagattcaacatttcctgtgcctctattccctttt 3772
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  146    AATATTGAAAAGGAGAGATGAGTATTCACATTTCCGTGTCGCCCTTATTCCTTTT 205
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  3773  ttgggcattttgccttctctgtttttgtcaccgaaacgtggtgaaagtaaatg 3832
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  206    TTGCGCATTTTGGCTTCTCTGTTTGTGTCACCCAGAACGCTGTTGAAAGTAAAGATG 265
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  3833  ctgaagatcagtttggttcagagtggtttacatcgaactgagatcgaacgggtgaaga 3892
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  266    CTGAGATCAGTGTGGGTGCACGAGTGGGTACATCGAATCGATCTCAACACGCGTAAGA 325
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  3893  tctttgagagttttccgcccgaagaaacgttttccaatgatgagcacttttaaaagtctgc 3952
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  326    TCCTTGACAGTTTCGCCCGGAGAACGCTTTCCAAATGATGACGACTTTTAAAGTTCGCG 385
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Qy  3953  tatgtggcgcgtattatcccgattattgacgcgggcaagagcaaacactcggtccgcgatac 4012
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  386    TATGTGGCGCGGTATTATCCCGTATTACGCGCGGCAAGAGCAACTCGGTCCGCCGATAC 445
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4013  actatttcagaatgacttgggttgagtactaccagtcacagaaagcatcttaagatg 4072
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  446    ACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCATCTTACGGATG 505
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4073  gcatgacagtgaagaattatcagtcgtgccataaccatgagtgataaactgagg-cc 4131
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  506    GCATCACAGTAGAGAAATATTCAGTGTGCTGCCATAACCATGAGTGATACACTGCGGCC 565
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4132  aacttactcttgacaacgacgagggagacgaggaagagcttaacgcgttttttgcacaacatg 4191
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  566    AACTTACTTCTGACACGATCGGAGGACCGAAGAGGACTAACCGCTTTTGTGCACACATG 625
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4192  ggggacatgataactcgccttgatcgttgggaacgagcagtgataaagcagcacaacac 4251
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  626    GGGGATCATGTAACTCGCCCTTGTCTGGGAACCGGAGCTGAATGAAGCCATACCAAC 685
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4252  gacgagcgtgacacccacgacgctgtgcaatggcaacaacgttgcgcaaacatttaact 4311
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  686    GACGAGCGTGACACCACGATGCTGTAGCAATGGCAACACCTTGGCACAACATTTAACT 745
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4312  ggcgaactacttactctgcttcccggaacaaattaatgactgagtgagggcgataaa 4371
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  746    GCGGACTACTTACTCTAGCTTCCGGCAACAATAATAGACTGGATGGRGCGGATAAA 805
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4372  gttcagagacacactctgctgctcccttcggtggtggtggttattgtctgataaatot 4431
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  806    GWTGAGGACCACTTCTGGCTCGGCCCTTCCGGCTGGGKTTATTCTGCTATAWATCT 865
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4432  ggagcgggtgagcgtgggtctgcggtatcattgacgactgggcccagatggttaagccc 4491
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  866    GGRGCGGKGAGCGTGGKCTCGCGTATCATTTGCAGCCTGGGGCAGATGGTAAGCCC 925
Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy  4492  tcc 4494
Db      ||
Qy  926  ccc 928
Db      ||

RESULT      8
AL042026    AL042026      841 bp      mRNA      linear      EST 29-FEB-2000
LOCUS       DKEZp434E111_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION  DKEZp434E111 5', mRNA sequence.
ACCESSION   AL042026
VERSION     AL042026.1 GI:5421372
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 841)
AUTHORS     Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE       EST (Poustka, et al.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Poustka A.J.
            Department Lehrach
            Max-Planck-Institute for Molecular Genetics
            Ihnestrasse 73, 14195 Berlin, Germany
            Tel: +49-30-84131623
            Fax: +49-30-84131128
            Email: poustka@mpg-berlin-dahlem.mpg.de
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No sl sequence available.
            This clone (DKEZp434E111) is available at the RZPD in Berlin.

```

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1. .841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFp434E111"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 226 a 192 c 202 g 221 t
ORIGIN

Query Match 14.6%; Score 841; DB 9; Length 841;
Best Local Similarity 100.0%; Pred. No. 4.2e-168; Mismatches 0; Indels 0; Gaps 0;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3639 tatttggtttatttcttaataacattcaaaatgatgtccgtcatgagacaataaacctg 3698
|||||
Db 1 TATTGTGTTATTTCTAAATACATTCAATATGTATCGCTCATGAGACAATAACCTG 60
QY 3699 ataaatgcttcaataatattgaaaaggagagatgatgtattcaacatttcggtcgc 3758
|||||
Db 61 ATAAATGCTTCAATAATATTGAAAAGGAAGATGATGAGTATTCAACATTTCCGTGTCG 120
QY 3759 cctattcccttttttgcgcatttgccttcctgttttgcctcaccagaaacgctggt 3818
|||||
Db 121 CCTATTCCCTTTTTCGGGCATTTTGCCTTCTCTGTTTTTTCCTCACCAGAAACGCTGGT 180
QY 3819 gaaagtaaaagatgctgaagatcagttgggtgcacgagtggtttacatcgaactggtatct 3878
|||||
Db 181 GAAAGTAAAGATGCTGAAGATCAGTTGGTGCACGAGTGGTTACATCGAATCGGATCT 240
QY 3879 caacagcggttaagatccttgagagatttgcgcccgcgaagaacgttttccaatgatgagcac 3938
|||||
Db 241 CAACAGCGGTAAAGATCCTTGAGAGTTTTTCGCCCGGAAGACGTTTTTCCAAATGATGAGCAC 300
QY 3939 ttttaagttctgtatgtgcccgttattatccgtattgacgcggggaagagcaact 3998
|||||
Db 301 TTTTAAAGTTCTGTTATGTGGCGGGTATTATCCCGTATTGACGCGGGCAAGAGCAACT 360
QY 3999 cggtcgccgcatacactattctcagaatgacttggttgagtactaccagtcacagaaaa 4058
|||||
Db 361 CGGTCGCCGATACATATTCTCAGAATGACTGGTTGAGTACTACCAGTCCACAGAAAA 420
QY 4059 gcatcttcagtgatggatgacagtaagagaattatgcagtgctgcataaccatgagtga 4118
|||||
Db 421 GCATCTACGGATGGGATGACATAGAGAAATTATGCAGTGTGCCATACCATGAGTGA 480
QY 4119 taacactgcggccaaacttactctgacaacgactcgaggaccggaagagctaacccgttt 4178
|||||
Db 481 TAACACTGCGGCCAACTTACTTCTGACACGATCGGAGGACCGAAGAGCTTAACCGCTTT 540
QY 4179 ttgcacaacatggggatcatatgaactgccttgatcggttggaaccggagctgaatga 4238
|||||
Db 541 TTTGCAACAATGGGGATCATGTAACATCGCCCTGTGCTGGGAACCGGAGCTGAATGA 600
QY 4239 agccataccaaacgagcggtgacaccagatgctgttagcaatggcaacacgttgcg 4298
|||||
Db 601 AGCCATACCAACGACGAGGCTGACACACGATGCCGTGTAGCAATGGCAACACGTTGCG 660
QY 4299 caaacttaactggcgaaacttactctactctagcttcccgcgaacaattaatagactggat 4358
|||||
Db 661 CAAACTATTAACTGGGCACTTACTTAGCTTCCCGGCAACATTAATACACTGGAT 720
QY 4359 ggaggcggtataaagttgcaggaccacttctgcgtcggccctcccggtcggctggtttat 4418
|||||
Db 721 GGAGGGGGATAAAGTTGCAGAGACCACTTCTGCGCTGGGCCCTTCCGCGTGGCTGTTAT 780
QY 4419 tgctgataaactcggagccggtgagcgtgggtctcgcggtatcatctgcagcactggggcc 4478
|||||

Db 781 TGTGTAAATCTGGAGCCGCTGAGCGTGTCTCGCGTATCATTCAGCACTGGGGCC 840
QY 4479 a 4479
Db 841 A 841
RESULT 9
BM438950
LOCUS
DEFINITION
BM438950
ACCESSION
VERSION
BM438950
KEYWORDS
EST.
SOURCE
ORGANISM
channel catfish.
Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
REFERENCE
1 (bases 1 to 854)
AUTHORS
Feng, J., Kucuktas H., Kocabas, A., Li, P. and Liu, Z.
TITLE
Transcriptome of channel catfish (Ictalurus punctatus): Initial
analysis of expressed sequence tags from the liver
JOURNAL
Unpublished (2002)
COMMENT
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse
FEATURES
source
1. .854
/organism="Ictalurus punctatus"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"
BASE COUNT 226 a 198 c 221 g 209 t
ORIGIN
Query Match 14.3%; Score 824.6; DB 10; Length 854;
Best Local Similarity 99.3%; Pred. No. 1.3e-164; Mismatches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 3778 gcatcttgccttctgttttgcacccagaaacgctggtgaaagttaaagatgctgaa 3837
|||||
Db 1 GCATTTTGCTTCTGTTTGTCTACCCACATACGCTGGTGAAGATGAAAGATGCTGAA 60
QY 3838 gatcagttgggtgcacagtggtg-ttaccatcgaactggatctcaacagcgggtaagatcct 3896
|||||
Db 61 GATCAGTTGGGTGCACGAGTGGGCTTACATCGAATCGATCTCAACAGCGGTAAATCCT 120
QY 3897 tgagagtttgcggcgaagaacgcttttccaatgatgagcaacttttaagttctgtctatg 3956
|||||
Db 121 TGAGAGTTTTCGCCCGGGAAGACGTTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATG 180
QY 3957 tggcgcgggtattatcccgattgacccgggcaagagcaactcggtcgccgcatacacta 4016
|||||
Db 181 TGCGCGGTATTATCCGGTATTGACCGCGGCAAGACCAACTCGGTGCGCGCATACACTA 240
QY 4017 ttctcagaatgacttggttgtagtactcaccagtcacagaaaaagactcttaccgtagtgcac 4076
|||||
Db 241 TTCTCAGATGACTTGGTTGAGTACTCACCAGTCACAGAAAGACATCTTACGATGGCAT 300
QY 4077 gacagttaagagaattatgcagtgctgctccataaaccatgagtgataaacactcgccgaactt 4136
|||||

```
Db 301 CACAGTAAGAGCAATTATGCTAGCTGCTGCCATAACCATGAGTGATTAACACTGCGGCCAACTT 360
Qy 4137 acttctgacaacgatcgagagacgaagagactaacgcgtttttttgacacacatggggga 4196
Db 361 ACTTCTGACAAAGATCGGAGAGACGGAA-GAGCTAACCGCTTTTGTGCACAAACATGGGGGA 419
Qy 4197 tcatgtaactcgcttgatcgcttgggaacggagctgaatgaagcaccacaaacgacga 4256
Db 420 TCATGTAACTCGCCTTGATCGCTTGGAAACCGGAGCTGNAATGAAGCCATACCAACGACGA 479
Qy 4257 gctgacacacacgatgctgttagcaatggcaacaacgcttgccaactatttaacttggcga 4316
Db 480 CGGTGACACACACGATGCTGTAGCAATGGCAACACGTTGCCAAACATATTAACTTGGCGA 539
Qy 4317 actacttacttaacttcccggaacaattaatagacttgatggagcgagataaagtgc 4376
Db 540 ACTACTTACTTACTTCCCGGCAACAATTAATAGACTGGATGGAGCGGATTAAGTTTGC 599
Qy 4377 aggaccacttgcgctcgccctccggctggctgggtttatttgcgtgataaactctggagc 4436
Db 600 AGGACCACCTTCTCGCGCTTCCGGCTGGCTGGTGTATTGCTGATTAATCTGGAGC 659
Qy 4437 cggtagcgtgggtctcgcggtatcattgcagcactggggccagatgtaagccctcccg 4496
Db 660 CGGTGAGCGTGGGTCTCGCGCTATCATTGCAGCACTGGGGCCAGATGTAAGCCCTCCCG 719
Qy 4497 tatcgtagtattctacacagcggggagtcaggcaactatgatgaacaaatagacagat 4556
Db 720 TATCGTAGTATTATACAGCAGCGGGAGTCAGGCAACTATGGATGAACGAGATAGACAGAT 779
Qy 4557 cgcgtgagatagtcctcactgattagcattgtaactgacacacacgaagttactcata 4616
Db 780 CGCTGAGATAGTGCCTCAGTATTAAGCATTTGGTAAGTGTGTCAGACCAAGTTTACTCAT 839
Qy 4617 tatacttagattga 4631
Db 840 TATGCTTGAGATTGA 854

RESULT 10
AU081044/c
LOCUS
DEFINITION
hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KBI,
mRNA sequence.
ACCESSION
AU081044
VERSION
AU081044.1 GI:6431392
KEYWORDS
EST.
SOURCE
rainbow trout.
ORGANISM
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE
1. (bases 1 to 1163)
Kono,T., Sakai,M. and Lapetra,S.E.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
CONTACT: Masahiro Sakai
Miyazaki University
Faculty of Agriculture
1-1 nishi gakuenkubandai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
Location/Qualifiers
1. .1163
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KBI"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"
```

```
BASE COUNT 272 a 307 c 298 g 286 t
ORIGIN
Query Match 13.6%; Score 782.4; DB 9; Length 1163;
Best Local Similarity 99.7%; Pred. No. 1.4e-155;
Matches 794; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 4972 ctgcctctgctaactctgtttaccagtgctgctgcagtgccgataaagtctgcttacc 5031
Db 1163 CTCGCTCTGCTAAATCCTCTGTACTCAGTGCTGCTGCCAGTAAAGTGTGCTGTACTAC 1104
Qy 5032 ggggttgactcaagcagatgttaccgataaagcgacgagcgtcggtcaacgggggt 5091
Db 1103 GGGTGGAGCTCAAGACGATAGTTACCGGATAAAGCGCAGCGTGGGCTGAAC-GGGGT 1045
Qy 5092 tcgtgacacacagccagcttgagcgaacgactacacggaactgagataacctacagcgt 5151
Db 1044 TCGTGACACAGCCAGCTTGGAGCGAAGCACTACACCGAACTCAGATACCTACAGCGT 985
Qy 5152 gagctatgagaaaagcgccacgcttccccgaaggagaaaaggaggtatccggttaagc 5211
Db 984 GAGCTATGAGAAAAGCGCCACGCTTCCCGAAGGGAGAGAGGATATCCGGTAAGC 925
Qy 5212 ggcagggctcggaacagagagcgacgagggaggttccaggggggaaacgcctggtatt 5271
Db 924 GGCAGGGTCGGAACAGAGAGAGCGCAGAGGGAGCTTCCAGGGGGAACGCCCTGTATCTT 865
Qy 5272 tatagtcctgctcggttttcgcacacctctgacctgagcgtcgatgttttctgactcgta 5331
Db 864 TATAGTCTCTGTGGGTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATCTCGTCA 805
Qy 5332 gggggggcgagcctatggaaaaaacccagcaacgcgccctttttacgggttccctgacct 5391
Db 804 GGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGCGGCTTTTACGGTTCTCGGCTTT 745
Qy 5392 tgcctggccttttgcacatgtttcttctcggttcttcctcggttaccctcgattctgtgga 5451
Db 744 TGCTGGCCTTTTGCTCACATGTTCTTCTCGGTTATCCCTGTATCTGTGGATAACCGT 685
Qy 5452 attaccgctttgagtgagctgataccgctcgccagccgacgacgagcgagcgagcgag 5511
Db 684 ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCGCAACCGAGCGGAG 625
Qy 5512 tcaagtgcgaggaagcggagagcgcccaatacgcgaacgcctctcccgcgcttgg 5571
Db 624 TCAGTGAGCGAGGAAGCGGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGGTTGG 565
Qy 5572 ccgattcattatgcagctggcgacgagaggtttcccgactggaagcgaggcgagtgagcg 5631
Db 564 CCGATTCAATTATGCGAGCTGGCAGCAGAGGTTTCCCGACTGGAAAGCGGCGAGTGGCGC 505
Qy 5632 aacgcaattaatgtagttagctcactcattagcaccacgagctttacactttatgctt 5691
Db 504 AACGCAATTATGTGAGTTAGCTCACTCATTAGGCAACCGCAGCTTTACACTTTATGCTT 445
Qy 5692 ccggctcgtatgttctgtggaattgtgagcggataacaatttcacacagggaacgctat 5751
Db 444 CCGGCTCGTATGTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAACAACGCTAT 385
Qy 5752 gaccatgattacgca 5767
Db 384 GACCATGATTACGCCA 369

RESULT 11
AU081449/c
LOCUS
DEFINITION
4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION
AU081449
VERSION
AU081449.1 GI:6929329
KEYWORDS
EST.
```


SOURCE ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE	COMMENT
African malaria mosquito.	Anopheles gambiae			
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae				
: Anopheles.				
1 (bases 1 to 800)				
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.				
Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)				
20300950				
Contact: Dimopoulos G				
Fotis C. Kafatos laboratory				
European Molecular Biology Laboratory				
Meyerohofstrasse 1, 69117 Heidelberg, Germany.				
Location/Qualifiers				
1. 800				
/organism="Anopheles gambiae"				
/strain="4A r/r"				
/db_xref="taxon:7165"				
/clone="4A3A-P405"				
/clone_lib="Anopheles gambiae immune competent 4A3A"				
/cell_line="immune competent 4A3A"				
/lab_host="E. coli DH10b"				
/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."				
BASE COUNT	203 a	206 c	198 g	193 t
ORIGIN				
Query Match	13.5%	Score 777.8;	DB 9;	Length 800;
Best Local Similarity	99.1%	Pred. No. 1.1e-154;		
Matches 782;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy 4620	actttagattgattaaaaacttcatttttaattcaaaaggatctaggtgaagatccctttt	4679		
Db 800	ACTTTAGATTGATTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAGATCCCTTT	741		
Qy 4680	tgataatctatgacaaaaatcccttaacgtgagtttcttcacactgagcgtagacccc	4739		
Db 740	TGATAATCTCATGACCAAAATCCCTTAAACGTGAGTTTTCGTTCCTACTGAGCGTCAGACCC	681		
Qy 4740	cgtagaagaatcaaaagatctcttgaagatccctttttctgcgttaactctgctt	4799		
Db 680	CGTAGAAGAATCAAAAGATCTCTTGAGATCCCTTTTTCTGCGCGTAATCTGCTGCTT	621		
Qy 4800	gcaaacaaaaaacccgcgtaccagcggtgtgtttgttgcggatcaagagctaccacac	4859		
Db 620	GCAACAAAAAACCCCGCTACCGCGGTGGTTGTTTGGCGATCAAGAGCTACCAAC	561		
Qy 4860	tcttttcggaaggttaactggcttcagcagcgagatcaacaaactgctcctctagt	4919		
Db 560	TCCTTTTCCGAAGGTAACTGGCTTTCAGCAGACGAGATACCAATACTGCTCTAGT	501		
Qy 4920	gtaccgtagtagggccaccactcaagaactctgtagcaccgcctacatcctcgtct	4979		
Db 500	GTACCGGTAGTTAGGCCACCACTTCAAGAACTGTGTAGCACCGCTACATACCTCGCTCT	441		
Qy 4980	gctaactcgtttacagtggtgctgctgcagtgggcgaataagctgctgtcttacccgggttga	5039		
Db 440	GCTAATCTGTTCACAGTGGCTGCTGCCAGTGGCGGATTAAGTTCGTGCTTACCGGGTTGA	381		
Qy 5040	ctcaagacgtagttaccggataaggcagcgatcggtggaacggggggtctgctgac	5099		


```
Matches 773; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4815 accgtaccagcgggtgtttgttccggatcaagagctaccacactcttttccgaagt 4874
|
|
|
Db 954 ACCGCTACCAGCGGTGGTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGT 895
|
|
|
Qy 4875 aactggttcagcagagcagatcaccaaaactctctcttagtctagcgcgttagtgg 4934
|
|
|
Db 894 AACTGGCTTCAGCAGAGCGGATACCAAAATACTGTTCTTCTAGTGTAGCGGTAGTTAGG 835
|
|
|
Qy 4935 ccaccactcaagaactctgtlagcaccgcctacatacctctctctctgctaactcttacc 4994
|
|
|
Db 834 CCACACTTCAAGAACTCTGTAGCACCCTTACATACCTCGCTCTGCTTAATCCCTGTACC 775
|
|
|
Qy 4995 agtggctctccagtgccgaataagctgtcttaccgggttggaactcaaacagatgtt 5054
|
|
|
Db 774 AGTGGCTCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTGGACTCAAGCATAGATT 715
|
|
|
Qy 5055 accgataagcgcagcgggtcggctgaacgggggttcgtgcacacagccagcttggga 5114
|
|
|
Db 714 ACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGA 655
|
|
|
Qy 5115 gcaagacacctacacccgaactgagatacctacagcgtgagctatgagaaagcgcacgt 5174
|
|
|
Db 654 CGGAACGACCTACACCGAACTGAGATACCTACAGCGCTGAGCTATGAGAAAGCCGACGCT 595
|
|
|
Qy 5175 tccgaaggggagaagcggagaggtatccggtgaagcgggggttcggaacagagagcgc 5234
|
|
|
Db 594 TCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCAGGGTCCGGAACAGGAGAGCG 535
|
|
|
Qy 5235 cagcaggagattccagggggaacgcctgtatcttattatagtcctcggggttcgcga 5294
|
|
|
Db 534 CACGAGGAGGCTTCAGGGGAAACGCGCTGTATCTTTATATGCTGCTGGGTTCGCGCA 475
|
|
|
Qy 5295 cctctgactgagcgtcatctttgtatgctcgtcagggggcgagccctatgaaaaa 5354
|
|
|
Db 474 CCTCTGACTTGAGCGTCGATTTTGTGATGCTGTCAGGGGGCGGAGCCCTATGAAAAA 415
|
|
|
Qy 5355 cgcagcaacgcgccttttacgggttcctggccttttgcctgttgcctttgtacatgtt 5414
|
|
|
Db 414 CGCCAGCAACGGCGCTTTTACGGTCTCCTGGCCCTTTTGTCTGGCCCTTTTGTCTACATGTT 355
|
|
|
Qy 5415 ctctcctggttaccctgattctgtgataacggtattaccgctttgagtagcga 5474
|
|
|
Db 354 CTTTCTCGGTATCCCTCGATTCTGTGGATAACCGGTATTACCGCTTTGAGTGAGCTGA 295
|
|
|
Qy 5475 taccgctcgcgcagccgaacagcagcagcagcagcagcagcagcagcagcagcagcagc 5534
|
|
|
Db 294 TACCGCTCGCGCAGCCGACGAGCGAGCGGAGCGAGTTCAGTGACGAGGAGCGGAAGA 235
|
|
|
Qy 5535 gcgcaccaatagcaaacgcctctccccgcgcgttgccgcgattcattaatgcag 5588
|
|
|
Db 234 CGGCCAATAGCAAAACCGCCTCTCCCGCGGCTTGGCGGATTCAATATGCAG 181
|
|
|

RESULT 13
BI489087/c
LOCUS
DEFINITION BI489087 990 bp mRNA linear EST 28-AUG-2001
603021122R1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 3',
mRNA sequence.
ACCESSION BI489087
VERSION BI489087.1 GI:15328315
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 990)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabp-r@mail.nih.gov
```

```
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11479 row: 1 column: 06
High quality sequence stop: 803.

FEATURES
    source
        1..990
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5191829"
                /clone_lib="NIH_MGC_114"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
                Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
                male brains, age range 23-27 yo. Library is oligo-dT
                primed and directionally cloned (EcoRV site is destroyed
                upon cloning). Average insert size 1.5 kb, insert size
                range 1-3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 019. Note:
                this is a NIH_MGC Library."
                BASE COUNT 174 a 316 c 318 g 182 t
                ORIGIN

Query Match 13.3%; Score 764.2; DB 10; Length 990;
Best Local Similarity 96.0%; Pred. No. 9.5e-152;
Matches 869; Conservative 0; Mismatches 28; Indels 8; Gaps 8;

Qy 1570 ctgtccccccagcggggcctgttccacagggcgtgctcagcagcgtgcccccaat 1629
|
|
|
Db 898 CTCTGTTCCCGCAACCGGGCAGCTGTCCACAGGCCCTGTCTCCAAACGGTCCCCCAAG 839
|
|
|
Qy 1630 ggaacctgggcaacgggtggcagggagggccctcgcagggccacgcagctggccac 1689
|
|
|
Db 838 GGA-CCTGGGCACCTGGGCCATGGAAAGGCCCGCTCGCA-GGCCACGCAGCTGGCCAC 781
|
|
|
Qy 1690 ctgtggcgtgtctccaggggcagcgtgtgggaatgaca-cagagcgtgtagctgct 1748
|
|
|
Db 780 CTGTG-TGGTGTCTTCCCA-GCGGCACCTGTGGGAATGACACACAGAGCTGTGCTGCT 723
|
|
|
Qy 1749 tcggacacagcagcagcagcgtctgtgtaaccacgaatggcagcgtgctcagaagaag 1808
|
|
|
Db 722 TC-GACACGACGCGCAGGTCTCTGTTGACCCAGCAATGGCAGCTGCTGCTCAAGAAAG 564
|
|
|
Qy 1809 cgtctccgggttctctcgtcgtgtgtgtagatggagacttctcagtgacacccaga 1868
|
|
|
Db 663 CGCTTCGGGTCTCTCTGCTGCTGTGTGTAGATGGAGACTTCTCTCAGTGACACCCAGA 604
|
|
|
Qy 1869 ggcctcctacacggggagagacttccacggcctgcagcagcgtgctgggtgtgtagaaga 1928
|
|
|
Db 603 GGCCCTCATCAACCGGGGAGACTTCCACGGCTCGCAGGTGCTGGTGGGT-TCGTGAAGA 545
|
|
|
Qy 1929 tgaggcctgctatttcttggttaccggggcccgagcgttcacgaagaacagcagtcct 1988
|
|
|
Db 544 TGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCAGGCTTCACCAAGACACAGAGTCTCT 485
|
|
|
Qy 1989 cctacgcggggcgcaggttctcgtggcggtgtcggttcgggggttccccagggttaagtac 2048
|
|
|
Db 484 CATACGCGGCTCGAGTTCCTGGCGGGGTGGGGTTCGGGGTTCGCCAGGTAAAGTACCT 425
|
|
|
Qy 2049 ggcacgcagcgtgtggttcctgcattacacagactggtgcctcccgagacccggcagc 2108
|
|
|
Db 424 GGCACCCGAGGCTGTGGTCTCATTTACACAGACTGGCTGTCATCCGAGGACCCGCGACG 365
|
|
|
Qy 2109 cctgaggagggccttgagcagatgtgtgggcagcacacaatgctggtgcccgtggccca 2168
|
|
|
Db 364 CCTGAGGAGGCGCCCTGAGCGATGTGTGGGGCCACCAATGTCGTGTCGCCCGTGGCCCA 305
|
|
|
```

QY 2169 gcttgcgggagctggtccaggggtcccggtctacgctacgtcttggagacacg 2228
 Db 304 GCTGGCTGGCGACTGGCTGCCAGGGTCCCGGGTCTACGCTACGCTTTTAAACACCG 245
 QY 2229 tgcttcacgctctctgcccctgtggatgggggtgcccacggtctacgagatcgatt 2288
 Db 244 TGTTCACGGCTCTCTGGCCCTGTGGATGGGGTGGCCACGGGTACGAGATCGAGTT 185
 QY 2289 catctttggatccccctggacccctctcgaactacacggcagagagaaatcttcgc 2348
 Db 184 CATCTTTGGGAT-CCGCTGGACCCCTCTCGAACTACAGGACAGAGAGAAATCTTCGC 126
 QY 2349 ccagcactgactgactactgggccaacttggccacacaggggatcccaatgagcccg 2408
 Db 125 CCAGCGACTGATGCGATCTGGGCCAATTTGCCCGACAGGGGATCCCAATGAGCCCCG 66
 QY 2409 agaccacaaagccccaatggcccctgtacacggcggggtccagcagtagcttactct 2468
 Db 65 AGACCCCAAGGCCCAATAGGCCCGCTACAGCGGGGGGTACGAGTACGTTAGTCT 6
 QY 2469 ggacc 2473
 Db 5 GGACC 1

RESULT 14
 BI488594
 LOCUS 603021122F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 5',
 DEFINITION mRNA sequence.
 ACCESSION BI488594
 VERSION BI488594.1 GI:15327822
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 931)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1479 row: 1 column: 06
 High quality sequence stop: 771.
 Location/Qualifiers
 1..931
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:5191829"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dr
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."
 BASE COUNT 152 a 294 c 306 g 179 t
 ORIGIN

Best Local Similarity 96.9%; Pred. No. 8e-149;
 Matches 882; Conservative 0; Mismatches 16; Indels 12; Gaps 11;
 QY 967 cgtggggccggcgtcggggcatcctcgaagacccccgggggctctctgtcttgc 1026
 Db 1 CGTGGGGCGCGCTGCGGGGCAATTCGCTGAAGACCCCCGGGGGCTCTCTCTGCTTTC 60
 QY 1027 ctgggcatcccttttggagaccacccatgggaccccgctgcttttgcacacggagccc 1086
 Db 61 CTGGGATCCCTTTTGGGAGCCACCATGGGACCCCGCTCTTCTGACCGGAGGCC 120
 QY 1087 aagcagccttggtcaggggtgtagcgtacacacctccagagtgctgtaccacat 1146
 Db 121 AAGCAGCCTTGGTCAAGGGGTGGTAGACGCTACAACCTTCAGAGTGTCTGCTACCAATAT 180
 QY 1147 gtgacacacctalaccacaggttttggggcaccgagatgtggaaacccccacgctgagctg 1206
 Db 181 GTGGACACCTATACCCAGGTTTGGAGGACACCGAGATGTGGAAACCCCAACCGTGAAGCTG 240
 QY 1207 agcggagactgcctgtacctcaactgttggaacacctacccccggcctacatccccacc 1266
 Db 241 AGCGAGGACTGCCTGTACCTCAACGTGTGACACCATACCCCGGCTACATCCCCCACC 300
 QY 1267 cctgtcctcgtctggtatctatgggggtgcttctacagtggggctcctcctctg-gacgt 1325
 Db 301 CCTGTCTCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCTTGTGACGT 360
 QY 1326 gtacgatggccgcttcttgggtacagccgagagactgtgctggtgtccatgaactaccg 1385
 Db 361 GTACGATGGCGCTTCTTGGTACAGCCGACAGAGACTGTGCTGGTGTCTTCACTGAACCTACCG 420
 QY 1386 ggtgggagcctttggcttctcctgcccgtccggggagccgagagccccggggcaatgtggg 1445
 Db 421 GGTGGAGAGCTTTGGCTTCTGTGCCCTCGCGGGAGCCGAGAGGCCCGCGGCAATGTGGG 480
 QY 1446 tctcctggtacagaggtcgtccctgagtggtgagggagacagctggagccttcggggg 1505
 Db 481 TCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGACAGGAAACGCTGGCAGCTTCGGGGG 540
 QY 1506 tgaccgcacacatcagtcagcgtgttggggagagcgcggagccgctcggtgggcatgca 1565
 Db 541 TGACCCGACATCAGTCACGCTGTTTGGGGAGAGCGCGGGAGCGCCCTCGTGGCATGCA 600
 QY 1566 cctgtcttcccccgcagcccggtg-cctgttccacagggcgctgtctgcagagc-ggtgcc 1623
 Db 601 CTTGCTGTCCCGCCAGCCGCGGGCCCTGTTCACAGGGCCGTGCTGACAGAGCGGTGCC 660
 QY 1624 cccaatggacccctggcaccggtgggcatgggagagccgctgcagggccac--gcagc 1681
 Db 661 CCAATGGACCCCTTGGCCACCGGTGGGCAATGGGAGAGGCCCGCTGCGCAGGCCACAGCGT 720
 QY 1682 tggccacaccttggggctgt-cctccagggcgctcctgggtgggaatgacacagagct-ggt 1739
 Db 721 TGGCCACCTTGTGGGCTGTCTCCCTCCAGCGGCGACCTGGTGGGAATGACACAGAGCTGGGT 780
 QY 1740 agcctgcc-ttcggacacagacc-agcgcaggttctcgtgtggaacacgaatggcagctgctg 1797
 Db 781 AGCTGCTTTTCGGACACACCAAGCGAGTCTGCTGGTGAACACCAATGGCAGCTGCTG 840
 QY 1798 cctcaagaagcgtcttcggttctcctcgtgctgctgtgtagatgagacttccctcagt 1857
 Db 841 -CTCAAGAAGCGTCTTCCGG-TCCTCTCTGTCGCTGT-GTCCATTTGCGACTTCTCTCAGG 897
 QY 1858 gacacccccg 1867
 Db 898 AAACCCAGAG 907

RESULT 15
 BG923768
 LOCUS 602825893F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954486 5',
 DEFINITION mRNA sequence.

Query Match 13.0%; Score 750.4; DB 10; Length 931;

ACCESSION BC923768
 VERSION BC923768.1 GI:14304244
 KEYWORDS EST.
 SOURCE house mouse
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 832)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9abs-r@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM10915 row: j column: 23
 High quality sequence start: 32
 High quality sequence stop: 823.
 Location/Qualifiers
 1..832
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4954486"
 /clone_lib="NCI CGAP Mam6"
 /sex="female_virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Sali;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"
 BASE COUNT 201 a 209 c 196 g 226 t
 ORIGIN
 Query Match 12.98; Score 743.6; DB 10; Length 832;
 Best Local Similarity 99.2%; Pred. No. 2.2e-147;
 Matches 768; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
 Qy 3904 ttctgcccgaagacgctttccaatgatgacgacttttaagtctgctgctggtggcg 3963
 Db 805 TTTGCCCGGGAAGACGTTTCCATGATGAGCAC-TTTAAGTTCGCTATGCGCGG 747
 Qy 3964 gtattatocgattagcgcgggcaagacactcgg-tcgccgcatacacactatttca 4022
 Db 746 GTATTATCCCGTATTGACGCGGGCAAGACAACTCGGTTCGCGCATACACTATTCTCA 687
 Qy 4023 gaatgacttggttagtactaccagtcacagaaagcattcttacggatgcatgacagt 4082
 Db 686 GAATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCATCTTACGGATGGCATGACAGT 627
 Qy 4083 aaggaattatgcagtgctgcataaccatgagtgataaacactcggcgaacttactct 4142
 Db 626 AAGAGATTATGCACTGCTGCGCATACCATGAGTGATANCACTGGCGCACTTACTTCT 567
 Qy 4143 gacaacgatcgaggaacgaaggagctaacccgctttttgcacaacatgggggatcatgt 4202
 Db 566 GACAACGATCGGAGACCGAAGGAGCTAACCGCTTTTTCACAACTGGGGGATCATGT 507
 Qy 4203 aactcgcttgatcgttgggaacggagctgaatgaagccataccaaacgacgagcgtga 4262
 Db 506 AACTCGCCTTATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGACGCTGA 447
 Qy 4263 caccacgatgctgtagcaatggcaaacacgcttgccgaactatttaactggcggaactact 4322
 Db 446 CACCACGATGCTGTAGCAATGGCANACGTTGGCCANAACTATTAACTGGCGAACTACT 387

Qy 4323 tactctagcttcccggaacaataatagactgagatggagcgggataaagtgcaggacc 4382
 Db 386 TACTCTAGCTTCCCGGCAACAATTAATAGACTGGAGCGGATAAAGTTGCAGGACC 327
 Qy 4383 acttctgctcgcgcccttccgctggttattgctgataaactgagccgggta 4442
 Db 326 ACTTCTGGGCTCGGCCCTTCCGGTGGCTGTTTATTGCTGATAAATCTGAGCGGTGA 267
 Qy 4443 gcgtgggtctcgcggtatcattgcagcactgggcccagatggtgaagccctcccgatcgt 4502
 Db 266 GCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGCCAGATGTAAGCCCTCCCGTATCGT 207
 Qy 4503 agttatctacacgacgaggagtcaggcaactatggatgaacgaataagacagatcgctga 4562
 Db 206 AGTTATCTACACGACGGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGA 147
 Qy 4563 gataggtgcctcactgattaaagcattgtaactgtcagaccaagttaactcatatatact 4622
 Db 146 GATAGGTGCTCACTGATTAAAGCATTTGTAAGTGTGACAGCCCAAGTTTACTCATATATACT 87
 Qy 4623 ttagattgatttaaaacttctattttaaatttaaaaggatctagggtgaagatcct 4676
 Db 86 TTAGATTGATTAAACCTCATTTTAAAGGATCTCGCGGAAGATCCT 33

Search completed: August 31, 2002, 15:25:56
 Job time: 6459 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:27:40 ; Search time 153.1 Seconds
(without alignments)
9252.570 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttcacgcctgcaggctc.....ctatgaccatgattacgcca 5767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2637.8	45.7	8854	3	US-09-053-549-1
C	2632.8	45.7	7387	4	US-09-238-356-28
C	2631.8	45.6	4283	1	US-08-343-401A-3
C	2631.8	45.6	4283	1	US-08-445-265A-1
C	2631.8	45.6	4283	3	US-08-990-442-1
C	2630	45.6	6295	2	US-08-659-206A-4
C	2628.2	45.6	4045	4	US-08-464-700-54
C	2628	45.6	4950	3	US-08-789-333F-58
C	2628	45.6	9639	4	US-09-147-208-26
C	2627	45.6	7566	2	US-08-232-016-23
C	2627	45.6	7639	2	US-08-232-016-22
12	2624.8	45.5	4713	4	US-09-194-285-7
13	2624.8	45.5	4724	4	US-09-194-285-8
14	2623.2	45.5	7560	4	US-08-844-274-20
15	2620.8	45.4	6350	2	US-08-385-335A-8
C	2616	45.4	11284	3	US-08-978-741-5
17	2579.8	44.7	3875	4	US-09-039-982A-21
18	2579.8	44.7	3875	4	US-09-039-641-21
19	2579.8	44.7	3875	4	US-09-039-762A-21
20	2579.8	44.7	3878	4	US-09-039-982A-27
21	2579.8	44.7	3878	4	US-09-039-641-27
22	2579.8	44.7	3878	4	US-09-039-762A-27
23	2579.8	44.7	3883	4	US-09-039-982A-30
24	2579.8	44.7	3883	4	US-09-039-641-30
25	2579.8	44.7	3883	4	US-09-039-762A-30
26	2579.8	44.7	3908	4	US-09-039-982A-24
27	2579.8	44.7	3908	4	US-09-039-641-24

28	2579.8	44.7	3908	4	US-09-039-762A-24	Sequence 24, Appl
29	2575.8	44.7	6111	4	US-09-538-414-9	Sequence 9, Appl
30	2545.6	44.1	3822	3	US-08-675-566-8	Sequence 8, Appl
31	2545.6	44.1	3861	3	US-08-675-566-11	Sequence 11, Appl
32	2545.6	44.1	3888	3	US-08-675-566-12	Sequence 12, Appl
33	2545.6	44.1	3955	3	US-08-675-566-10	Sequence 10, Appl
34	2545.6	44.1	4009	3	US-08-675-566-9	Sequence 9, Appl
35	2545.6	44.1	4503	3	US-08-675-566-7	Sequence 7, Appl
36	2544.6	44.1	5067	3	US-08-675-566-20	Sequence 20, Appl
37	2544.6	44.1	5147	3	US-08-675-566-24	Sequence 24, Appl
38	2544.6	44.1	5241	3	US-08-675-566-23	Sequence 23, Appl
39	2451.6	42.5	10306	3	US-08-716-351A-4	Sequence 4, Appl
40	2451.6	42.5	10970	3	US-08-716-351A-5	Sequence 5, Appl
C	2449.4	42.5	5639	5	PCT-US94-00658-1	Sequence 1, Appl
C	2437.4	42.3	4800	5	PCT-US94-07779-1	Sequence 1, Appl
C	2418	41.9	8225	4	US-08-793-618-1	Sequence 1, Appl
45	2416.4	41.9	19307	3	US-08-836-022A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-053-549-1
; Sequence 1, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,549
FILING DATE: 01-APR-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8854 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 11..3694
OTHER INFORMATION: /product= "hyFLIB protein"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8854
OTHER INFORMATION: /note= "Sequence of pCIB5520"
Patent No. 6121521
OTHER INFORMATION: containing coding sequence for hyFLIB protein"
US-09-053-549-1


```

: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: pWRG1630
: FEATURE:
: NAME/KEY: exon
: LOCATION: 713..721
: FEATURE:
: NAME/KEY: exon
: LOCATION: 981..1253
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(713..721, 981..1253)
: NAME/KEY: sig_peptide
: LOCATION: 713..1049
US-08-343-401A-3

Query Match 45.6%; Score 2631.8; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3133 aattcaactggcgcgtcgcttttacaacgctgctgactgggaaacccctggcggttaccacaactt 3192
Db 4265 AATTACATGGCGCGTGGTTTATCAACGTCGTCGACTGGGAAACCCCTGGCGTTTACCCAACCTT 4206

Qy 3193 aatcgcttcgagcaacatcccccttcgccagctggcgtggaatagcgaagggccgcacc 3252
Db 4205 AATGCCCTTGAGCAGCATCCCCCTTTCGCCAGCTGGCGCTGATAGCGAAGAGCGCCGCAACC 4146

Qy 3253 gatcgcccttccccacagcttgcgcagcctgaatggcggaatggcgctgatcggtatttt 3312
Db 4145 GATGCCCTTCCCAACAGTTGCGAGCTGAATGGCGAATGGCGCTGATGGCGTATTTT 4086

Qy 3313 ctccctacgcatctgctggtatttcacacgcacatagtgcaactctcagtaacatcgc 3372
Db 4085 CTCCTTACGATCTGTCGGGTATTTACACCGCATATGGTGACATCTCAGTACAAATCTGC 4026

Qy 3373 tctgatgcgcgatagtttaagccagcccgacaccccgcaacaccccgctgaagcgccctga 3432
Db 4025 TCTGATGCGCATAGTTAAGCGAGCCCGACACCGCCCAACACCGCTGACGCGCCCTGA 3966

Qy 3433 cgggcttctcgtctcccgcatcgcgttacagacaagctgtgacgcgtctccggagctgc 3492
Db 3965 CGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 3906

Qy 3493 atgtcagaggttttccacgcatcacgaaacgcgagacgaaagggcctcgatgata 3552
Db 3905 ATGTGTCAGAGGTTTTCACCGTCTATCACCGAAACCGCGAGACGAAAGGGCTCTGTGATA 3846

Qy 3553 cgcctattttataggttaagtcatgataataatggtttcttagacgtcaggtggcact 3612
Db 3845 CGCCTATTTTATAGTTAATGTCATGATATAATGGTTTCTTAGACGCTCAGGTGGCACT 3786

Qy 3613 ttctggggaaatgctgcggaaacccctatttggttattttcttaataacattcaaatatg 3672
Db 3785 TTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTTAATACATTCAAATATG 3726

Qy 3673 tatcgctcatgagacaataacccctgataaatgcttcaataatattgaaaggaagagt 3732
Db 3725 TATCGCTCATGAGACAATAACCCCTGATAATGCTTCAATATATGAAAGGAAGAGT 3666

Qy 3733 atgagatttcaacatttccgtgctgccttatcccttttttggcgcattttggccttcc 3792
Db 3665 ATGAGTATTAACATTTCCGTGTCGCCCTTATTCCTTTTTCGGGCATTTTGCCCTTCT 3606

Qy 3793 gttttgtcaccacgaacgctgggtgaagtaaaagatgctgaagatcagttgggtgca 3852
Db 3605 GTTTTGTCTACCCAGAACGCTGGTGAAGTAAAAAGATGCTGAAGATCAGTTGGGTGCA 3546

```

```

Qy 3853 cgaatgggttacatcgaactcgaatcctcaacagcggtaagatccttgagagttttccgccc 3912
Db 3545 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGTAAGATCCTTTGAGAGTTTTCGCCCC 3486

Qy 3913 gaagaacgttttcccaatgatgagcacttttaagttcttgctatgtggcgcgctattatcc 3972
Db 3485 GNAAGAACGTTTCCCAATGATGAGCACTTTTAAAGTTCGTATGTGGCGCGGTATTATCC 3426

Qy 3973 cgtattgacgcggggcaagagcaactcgtgcgccatcacactattctcagaatgacttg 4032
Db 3425 CGTATTGACGCGCGGCAAGAGCAACTCGGTCCGCCCATACACTATTCTCAGAATGACTTG 3366

Qy 4033 gttgagtactcaacagtcacagaaaagcatcttaacgattgagcatgacagtaagaaatga 4092
Db 3365 GTTGTAGTACTCACCGAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTA 3306

Qy 4093 tgcagtgctgccataaacatcagtgataacactgcgcgaacttacttctgacaacgac 4152
Db 3305 TGCAGTGTGCTCCNTAACCATGAGTGATACACTGCGGCCAACTTACTTCTGACACGATC 3246

Qy 4153 ggagaccggaaggagctaacccgctttttgcaacaacatgggggatacatgttaactcgctt 4212
Db 3245 GGAGCACCGAAGGAGCTAACCGCTTTTTCACACAACATGGGGGATCATGTAACTCGCCCTT 3186

Qy 4213 gatcgttgggaaaccgagcgaatgaagccataaccaaacgacgacgctgacaccacgatg 4272
Db 3185 GATCCTTGGGAAACCGAGCTGAATGAAGCCATACCAACGACGACGAGCTGACACCAAGTG 3126

Qy 4273 cctgtagcaatggcacaacgcttgcgcaaacctataaactgcggaacttacttacttagct 4332
Db 3125 CCGTAGCAATGCCAACAACGTTGGCGAANACTTATTAACTGGCGAATCTTACTCTAGCT 3066

Qy 4333 tcccggcaacaatataactgagtgagggcggaataaagttgcagggaacacttcttcgctg 4392
Db 3065 TCCCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCSC 3006

Qy 4393 tcggcccttcggcgctggctgttttctgataaacttgagacgagcggtgagcggtggtct 4452
Db 3005 TCGGCGCTTCGGCGTGGCTGGTGTATGCTGATAAATCTGGAGCGGTGAGCGTGGGTCT 2946

Qy 4453 cgcgtatcattgcagcactggggccagatggaagccctccctcatgtagttatctac 4512
Db 2945 CGCGTATCATTTGCAGCACTGGGGCCAGATGGTAGCCCTCCCGTATCGTAGTTATCTAC 2886

Qy 4513 acgacggggagtcaggcaactatggtgaacgaataagacagatcgcgtgagatgagtc 4572
Db 2885 ACGACGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGTGCC 2826

Qy 4573 tcaactgattaagcattggttaactgtcagacaagtttactcataataactttagattgat 4632
Db 2825 TCACGTGATTAAGCATTTGGTAAGTGTACAGACCAAGTTTACTCATATATATCTTTAGATTGAT 2766

Qy 4633 ttaaaactctcattttaattaaaaggatcctaggtgaagatcctttttgataatctcatg 4692
Db 2765 TTAACAATTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATTAATCTCATG 2706

Qy 4693 accaaaaatcccttaacgtgagtttctccactgagcgctgagaccccgtagaaaagatc 4752
Db 2705 ACCAAAAATCCCTTAACGTTAGTTCGTTCCACTGAGCGTCAGACCCCGCTAGAAAAAGATC 2646

Qy 4753 aaagatattcttgagatccctttttctgcgcgtaactgtctgtctgcaacaaaaaaa 4812
Db 2645 AAAGGATCTTCTTGAGATCTCTTTTTCGCGCGTAATCTGCTGTGTGCAACAAAAAA 2586

Qy 4813 ccacccgtaccagcggtgtgtttgttcgcggatcaagagctaccactctttttccgaag 4872
Db 2585 CCACCGCTACCAAGCGGTGGTTTGTGGCGGATCAAGAGCTACCAACTCTTTTTCGGAAG 2526

Qy 4873 gtaactgcttcagcagcgcatgataccaaataactgtcttcttagtgcgctagttga 4932
Db 2525 GTAACGTGGCTTCAGCAGAGCGCAGATACCAAACTACTGTCTTCTTAGTGAGCCGTAGTTA 2466

Qy 4933 ggcacacacttcaagaactctgtagacgcgcctacataactcgtctcgtctaatctgtta 4992

```

```
Db 2465 GGCCACCACCTCAAGAACTCTGTAGACACCGCTACATACCTCGCTCTGCTAATCTCTTA 2406
QY 4993 ccagtggctgctgcagtgccgataaactgctgtcttaccgggttgagactcaagacgtag 5052
Db 2405 CCAGTGGCTGCTGCCAGTGGGATAAGTCTGTCTTACCGGGTGGACTCAAGACGATAG 2346
QY 5053 ttaccggataaagcgagcggctcggtgacgaggggggttcctgacacacagccagcttg 5112
Db 2345 TTACCGGATAAAGCGCAGCGCTCGGGTGAACGGGGGTTCGTGCACACAGCCAGCTTG 2286
QY 5113 gaggcaagcctacacaggaactagatacctacagcgttagctatagagaagcgccacg 5172
Db 2285 GAGGGAACGACCTACACCGAACTGAGATACCTACAGCTGAGCTATGAGAAAGCGCCAG 2226
QY 5173 ctctccgaagggagaaagcgagcaggtatccggtgaagcgagggcgaggaagagagag 5232
Db 2225 CTTCCTCCGAAGGAAAGCGGACAGGTATCCGTAAGCGCAGGGTCGGAACAGGAGAG 2166
QY 5233 cgaacgagggagcttccagggggaacgcctggtatctttatagctcgtcggtttcgc 5292
Db 2165 CGCAGGAGGAGCTTCAGGGGGAACGCTGTATCTTTATAGTCTGTCTCGGGTTTCGC 2106
QY 5293 cactctgactgagcgtcgtatctttgtgctgctgcagggggggcgagccctatgaaa 5352
Db 2105 CACCTCTGACTTGAGCTGCGATTTTGTGATGCTGCTGAGGGGGGGGAGCCTATCGAAA 2046
QY 5353 aagcgagcaacggcgcttttacggttctgctgctgcttttgccttttgccttttgcctatg 5412
Db 2045 AAGCCGACGAAAGCGGCTTTTACGGTTCCTGGCCCTTTTGTGGCCCTTTTGTCTCATG 1986
QY 5413 ttcttctcgttattccccctgattctgtgtgataaacgctattaccgctttgagtgact 5472
Db 1985 TTCTTTCTGCTGCTATCCCTGATTTCTGTGGATAACCGTATTACCGGCTTTTGTAGTGACT 1926
QY 5473 gatacgtctgcgcagcgaacgagcagcagcagcagcagcagcagcagcagcagcagcag 5532
Db 1925 GATACGCTCGCCGACGCGAAGCGGAGCGGAGCGAGTCACTGAGCGAGGAAAGCGGNA 1866
QY 5533 gaggcgcccaatacgcaaacgcgctctcccgcgcttcccgcgcttcccgcgcttcccgcgct 5592
Db 1865 GAGCGCCCATACGCAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATGACAGCTGG 1806
QY 5593 cagcagaggttccccgactggaagcggcagtgagcagcagcagcagcagcagcagcagcag 5652
Db 1805 CAGCAGAGGTTTCCGACTGGAAGCGGCGAGTCACTGAGCGCAACGCAATTAAATGAGTTAG 1746
QY 5653 ctactcattagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 5712
Db 1745 CTACATATTAGGACCCCGAGGCTTTACACTTTATGCTTCCGGCTCGATGTTGTGTGGA 1686
QY 5713 attgtgagcggataaacaatttcacaggaagcagcagcagcagcagcagcagcagcag 5767
Db 1685 ATTTGAGCGGATAAACAATTTCACAGGAAACAGCTATGACCATGATTAGGCCA 1631
```

RESULT 4

```
US-08-445-265A-1/c
; Sequence 1, Application US/08445265A
; Patent No. 5697901
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Elof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,265A
; FILING DATE:
; CLASSIFICATION: 604
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 110229.91080
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Plasmid DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1250)
; US-08-445-265A-1
```

```
Query Match 45.6%; Score 2631.8; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3133 aattcactggcgctgttttacaacgctgctgactgggaaacccctggcttaccacactt 3192
Db 4265 AATTCATGCGCGCTGCTTTTACACGTCGTGACTGGGAAACCCCTGGCGTTACCAACTT 4206
QY 3193 aatgccttgcagcaatcccccttgcgcagctggcgtaataagcgaagagggccgcacc 3252
Db 4205 AATCGCGCTTCAGCACATCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCACC 4146
QY 3253 gatcgccctcccaacagttgcagcagcctgaatggcgaatggcgctgacggtattttt 3312
Db 4145 GATCGCGCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTT 4086
QY 3313 ctcttcagcactgtgtcggtattttcacacgcgcatatgggtgcactctcagtcacactgc 3372
Db 4085 CTCCTTACGCATCTGTGCGGTATTTTACACCGCATATGGTGCACTCTCAGTACAACTGC 4026
QY 3373 tctgatgcgcagctagttaagcagcagcagcagcagcagcagcagcagcagcagcagc 3432
Db 4025 TCTGATGCCGCATAGTTTAAAGCCAGCCCGACACCGCCCAACCCCGCTGACGCCCTGA 3966
QY 3433 cgagctgtctgtcccgccatccgcttcagacagcagcagcagcagcagcagcagcagc 3492
Db 3965 CGGGCTGTCTGTCCCGCATTCGCTTACAGCAAGCTGTGACCGTCTCGGGAGGTGC 3906
QY 3493 atgttcagaggttttcacgcgtcatcaccgaaacgcgcgagacgaaagggcctcgtgata 3552
Db 3905 ATGTGTCAGAGGTTTTCACCGTCATCACCCGAAACCGCCGAGAGCAAGAGGCCCTCGTGATA 3846
QY 3553 cgctattttataggttaagttaagttaagttaagttaagttaagttaagttaagttaagt 3612
Db 3845 CGCCTATTTTATAGGTTAATGTCATGATATAATGTTTCTTAGACGTCAGGTGGCAGT 3786
QY 3613 ttctgggaaatgtgcggaacccctattgttttatttttcttaataacacattcaaatg 3672
Db 3785 TTTCGGGAAATGTGCGGAAACCCCTATTGTATTATTTCTAAATACATTAATCAATATG 3726
QY 3673 tatccgctcatgagacaataaacctcgtataaaatgcttcaataatattgaaaaagagagt 3732
Db 3725 TATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGAGAGT 3666
QY 3733 atgagtattcaacatttcccgctgcgcacctattcccttttttgcgcgcaatttgccttcc 3792
```

3665 ATGAGTATTCACATTTCCGTCGCGCTTATTCCTTTTTCGCGCATTTTGCCCTTCT 3606
 3793 gttttgtcaccggaacgcgtggtgaaagttaaagatgctgaagatcagttggtgca 3852
 3605 GTTTTGTCTACCCAGAACGCTGCTGAAAGTAAAGATGCTGAAGATCAGTTGGTGCA 3546
 3853 cgaatggttaacatcgaactgagatctcaacacgcggtgaagatccttgagagatttcgccc 3912
 3545 CGATGGGTATACATCGAATCGAATCTCAACACGCGTAAAGATCTTGAGAGTTTTCGCCCC 3486
 3913 gaagaacgttttccaatgatgagcacttttaagttctctgctatggtgagcggtattatcc 3972
 3485 GAAGAACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGCTGCGCGGTATTATCC 3426
 3973 cgtattgacgcccgggcaagagcaactcggctgcgcacatacactattctcagaatgacttg 4032
 3425 CGTATTGACGCGCGGCAAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTG 3366
 4033 gttgagtactcaacagtcacagaaagcatcttaacggtgcatgagcaagtaagaaatta 4092
 3365 GTTGAAGTACTACCACTACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTA 3306
 4093 tgcagtgtctgcaataaccatgagtgataaacactgcggccaacttaactctgacaacgac 4152
 3305 TGCAGTGTGCCATTAACCATGAGTGATTAACACTGGCGCCAACTTACTTCTGACACAGATC 3246
 4153 ggagaccgaagagactaacccgttttttgcacaacatggtgggacatgtaactgcgctt 4212
 3245 GGAGACCGAAGGAGTAAACCGCTTTTGTGCACAACATGGGGGATCATATACTCGCCTT 3186
 4213 gatcgttgggaacccgagctgaatgaagccatacacaacacgacgagcgtagaacaccagtg 4272
 3185 GATCGTTGGGAACCGGAGTGTAATGAAGCCATACCAACGACGAGCGTGACACCAAGATG 3126
 4273 cctgtagcaatggcaaacgcttgcgcaaacattataactggtgcgaactactactctagct 4332
 3125 CCGTAGCAATGCCAACACGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAGCT 3066
 4333 tcccggaacaaatgaatagactggtgagggcggaataagttgcaggaccactcttgccg 4392
 3065 TCCCGCAACAAATTAATAGACTGGATGGAGGGGGTAAAGTTGCAGGACCACTTCTGCGC 3006
 4393 tcggccctccgctgagctggtttattgctgataaacttggaacccggtgagcgtagggtct 4452
 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAATCTGGAGCCGGTGAGCGTGGGTCT 2946
 4453 cgggtatcatctgcagcactggggccagatggtgaagccctccglatcgtagttatctac 4512
 2945 CGCGGTATCATTCAGCACTGGGGCCAGATGCTAAGCCCTCCCGTATCGTACTTATCTAC 2886
 4513 acgacggggagtcaggcaactatgtagaacgaaatagacagatcgctgagatagtgcc 4572
 2885 ACGAGGGGGAGTCAGGCAACTATGGATGAACAAATAGACAGATCGCTGAGATAGTGCC 2826
 4573 tcactgattaacagctggttaactgacagccagtttaactatataactttagattgat 4632
 2825 TCACGTATTAAAGCATTTGTTAACTGTGACACCAAGTTTACTATATATCTTTAGATTGAT 2766
 4633 taaaactctcatcttaattaaaaggatctaggtagaagatcctttttgataatcctcag 4692
 2765 TTAAACATTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATATCTCATG 2706
 4693 accaaatcccttaacgtgagtttctggtccactgagctcagagcccccagtagaaagatc 4752
 2705 ACCAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGCTAGAAAAAGATC 2646
 4753 aaaggatctcttgagatccttttttctgcgcgtaactgctgctgcaacaaacaaaaa 4812
 2645 AAAGGATCTCTTGAGATCCTTTTTCGCGTAATCTGCTGTGCAACAAAAA 2586
 4813 ccacgcctaccagcgtggtttgttgcgagatcaagagctaccactcttttccgaag 4872

2585 CCACCGCTACCAGGGGTGTTGTTTTCGCGGATCAAGAGCTACCAAACTCTTTTTCGGAAG 2526
 4873 gtaactggcttcagcagagcgagataccaaatactgtctctcttagtgcagcgtagtta 4932
 2525 GTAACGTGGCTTACAGAGAGCGCAGATACCAAACTACTGTCTCTTCTAGTGTAGCGTAGTTA 2466
 4933 ggcacacacttcaagaactctctagcacgcgcctacatacactcgtctctgtaatactctgta 4992
 2465 GGCACACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTTAATCTCTGTTA 2406
 4993 ccagtggctgcgcagcagtgccgataaagtcgtcttaccggggttgagactcaagacgatatg 5052
 2405 CCAGTGGCTGCTGCCAGTGGCGATAGTCTGCTCTTACCGGGTTGGACTCAACACGATAG 2346
 5053 ttaccggataaaggcgcagcgtcggtgaaacgggggttcgtgcacacagcccagcttg 5112
 2345 TTACCGGATAAAGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAGCCACGCTTG 2286
 5113 gagcaaacacatacaccgaactgagataccttacacgctgagcctatgagaagaagccacg 5172
 2285 GAGCGAACACCTTACACCACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACG 2226
 5173 ctcccggaagggagaaagcgacaggtatccggttaacggttaacggttcggaacagagag 5232
 2225 CTTCGGAAGGGAAGGCGGACAGGTATCCGTTAAGCGGCGAGGGTCSGAAACAGAGAG 2166
 5233 cgacagggagcttcacagggggaaacgcctggtatctttagctcctgcggggttcgc 5292
 2165 CGCACGAGGAGCTTCCAGGGGGAACGCTGCTATCTTTATAGTCTCTGTGCGGTTTCG 2106
 5293 cactctgactgagcgtcgattttgtgactgcgtaacgggggagcctatgagccttatggaaa 5352
 2105 CACCTCTGACTTGTAGCGTGCATTTTGTGATGCTCTCAGGGGGCGGAGCCCTATCGAAA 2046
 5353 aacgcagcaacgcggccttttaacggttccctggtcctttgtgcgcttttgcctcacatg 5412
 2045 AACGCCAGCAACGGCGCTTTTACGGTTCTTGGCCCTTTTGTGGCCTTTTGTGCTACATG 1986
 5413 ttcttctcgtctatccctgattctgttgataacccgtattaccgcctttgagtgagct 5472
 1985 TTCTTTCTCTGCTTATCCCTGATTTCTGTGATTAACCGTATTAACCGCTTTGAGTGAGCT 1926
 5473 gatacgcctccgcagcagcaacgcagcagcagcagcagcagcagcagcagcagcagcagc 5532
 1925 GATACCGCTCGCGCAGCGCAACGACACGAGCGCAGGTCAGTGAGCGGAGGAGCGGAA 1866
 5533 gagcgcccaatacgcgaacgcctctcccgcgctgtggcggttggcggttcaataatcagctgg 5592
 1865 GAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGCTTGGCCGATTCATTAAATGCAGCTGG 1806
 5593 cagcagaggtttccgcagctggaagcggcgagtgagcgcgaacgaatgaatgtgagttag 5652
 1805 CACGACAGGTTTCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG 1746
 5653 ctaactcattagggcaacccagcgttttaacatttatcttccggcgtgattgtgtgga 5712
 1745 CTCACATTAATAGCACCCCGAGGCTTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGA 1686
 5713 attgtagcggatacaaatctcacacaggaacacagcgtatgaccatgattacgcca 5767
 1685 ATTGTAGCGGATACAAATTTACACAGGNAACAGCTATGACCATGNTTACGCCA 1631

RESULT 5
 US-08-990-442-1/c
 ; Sequence 1, Application US/08990442
 ; Patent No. 6090790
 ; GENERAL INFORMATION:
 ; APPLICANT: Eriksson, Elof
 ; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Quarles & Brady

Db 5995 TTGTCGTCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTG 5936
QY 3498 tcagaggttttcaccgctcatcacccgaaacgcgcgagacgaaagggcctcgtgatcacgctt 3557
Db 5935 TCAGAGGTTTTTACCGGTCATCACGGAACCGCGAGACGAAAGGGCCTCGTGATACGCTT 5876
QY 3558 atttttataaggttaataatgataataatgggtttctttagacgtcaggtggcaactttcg 3617
Db 5875 ATTTTATATAGTTAATGTCATGATAATAAATGGTTTCTTAGACGCTAGGTGGCACTTTTCG 5816
QY 3618 ggaaatgtcgcggaaccctatttggttatttttcttaataatcattcaaatatgtatcc 3677
Db 5815 GGGAAATGTGCGGGAAACCCCTATTGTGTTATTTTCTAAATACATCAAAATGTATCC 5756
QY 3678 gctcatgagacaataaccctgataaatgtctcaataataattgaaaggaagatgatgag 3737
Db 5755 GCTCATGAGACAATAACCCGTGATAATGCTTCAATAATAATTGAAAGGAAGAGTATGAG 5696
QY 3738 taitcaaatcttcggtgtcgcccttattcccttttttttggcgcaattttgccttctgtttt 3797
Db 5695 TATTCAACATTTTCCGTGTCCGCTTATTTCCCTTTTTCGCGCATTTTGGCTTCCCTGTTTT 5636
QY 3798 tgcctcacccgaaacgctggtgaaagttaaagatgctggaagatcagttgggtgcacgagt 3857
Db 5635 TGCTACCCAGAAACGCTGCTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACAGT 5576
QY 3858 gggttacatgaaactggatctcaacagcggtaagatccttgagagtttttgcgcccgaaga 3917
Db 5575 GGGTTTACATCGAATGGATCTCAACAGCGGTAAAGTCTTGAGAGTTTTCGCCCGGAAGA 5516
QY 3918 acgttttccaatgatgagcaactttttaaagttctgctatgtggcgcggtattatcccgtat 3977
Db 5515 ACCTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTTATCCCGTAT 5456
QY 3978 tgacgcgggcaagagaaactcggctgcgcacatacacattcttcagpatcacttggttga 4037
Db 5455 TGACGGCGGCAGAGCACTCGGTGCGCCGCATACACTATTCAGAAATGACTTGGTTGA 5396
QY 4038 gtactcacccagtcacagaaagcatcttacggatggcgatgacagtagaagaattatgcag 4097
Db 5395 GTACTCACCAAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAAAGAAATTTATCGAG 5336
QY 4098 tgcgtccataaccatgagtgataaacactgcgcgaacttacttctgacaacgatcgagg 4157
Db 5335 TGTGTCATTAACCATGATGATTAACACTTGGCGCCAACTTACTTCTGACAAAGATTCGGAGG 5276
QY 4158 accgaaggagctaacgcgttttttgacaacatgggggatacatgtaaactgccttgatcg 4217
Db 5275 ACCGAAGGAGCTAACCCTTTTTCACAAACATGGGGGATCATGTAACTCGCCTTGATCG 5216
QY 4218 ttgggaacggagctgaatgaagccataaccaaacgacgagcgtgacacacgatgcctgt 4277
Db 5215 TTGGGAACCGGAGCTGAATGAAGCCATATCCAAACGACGAGCGTGACACCACGATGCCGT 5156
QY 4278 agcaatggcaaacagttgcgcaactattaaactggcgaaacttacttacttagcttcccg 4337
Db 5155 AGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTACTTACTAGCTTCCCG 5096
QY 4338 gcaacaattaatagactggatggagcggaataaagtgtgagggaccaacttctgcctcggc 4397
Db 5095 GCAACAATTAATAGCTGGATGAGGCGGATTAAGTTGCAGGACCACCTTCTGCGCTCGGC 5036
QY 4398 ccttcgggctgggtgtttattgtctataaattctgagcgggtgagcgtgggtctgcg 4457
Db 5035 CTTTCGGCTTGGCTGGTTTATTGCTGATAAATCTGGAAGCGGTTGAGCGGTGGGTCTCGCGG 4976
QY 4458 tatcattgagcaactggggccagatggtaagccctcccgatatcgtagttatcacagac 4517
Db 4975 TATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTTCCCGTATCGTAGTTATCTACAGCAC 4916
QY 4518 gggagtcaggcaactatggatgaacgaataatagacagatcgctgagataggttccctcaact 4577
Db 4915 GGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGTGGCTTCACT 4856

QY 4578 gattaagcattggtaactgtcagaccgaagtttactcatatatactttagatttataaa 4637
Db 4855 GATTAAAGCATTTGGTTAACTGT'CAGACCAAGTTTACTCATATATACTTTAGATTGATTAAA 4796
QY 4638 acttcatttttaattttaaaggatctagggtgaagatccttttttgataatctcatgaccaa 4697
Db 4795 ACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCATGACCAA 4736
QY 4698 aatcccttaacgtgagtttccgtccactgagcgtcagaccgcgtagaaaagatcaaaagg 4757
Db 4735 ANYCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAG 4676
QY 4758 atctcttgagatacctttttttctgctgctaactgctgtgcataacataacaaaaaacacc 4817
Db 4675 ATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGTGCTGTCATACTCGTCTGTGTACCAGT 4616
QY 4818 gctacacgaggttggtttggtttgcggatcaagagactaccaactctttttccgaaggtaac 4877
Db 4615 GCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGGAAGGTAAAC 4556
QY 4878 tggcttcagcagcagcgcagatacacaataactgtccttctagtgtagcgttagtgagcca 4937
Db 4555 TGCGTTCAAGAGCGGAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCA 4496
QY 4938 ccacttcaagaactctgttagcacccgctacatacctcgtctcgtctgctaatcctgtttaccagt 4997
Db 4495 CCAGTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCTGTGTACCAGT 4436
QY 4998 ggtcgtgcagtggtggatgaagtgcttcttacccgggttggactcaagagagatgttacc 5057
Db 4435 GGCTGCTGCGCAGTGGCGATTAAGTCTGCTTACCGGTTTGACTCAAGAGCATAGTTACC 4376
QY 5058 ggataaaggcgcagcgtcgcggtgaacgggggttcgtgcacacagcccagcttggagcg 5117
Db 4375 GGATAAGCGCAGCGGTCGGGCTGGAACGGGGGTTCTGTGCACACAGCCAGCTTGGAGCG 4316
QY 5118 aacgacctacacggaactgagataccttacacgctgagctatgagaaaagcgcacgcttcc 5177
Db 4315 AACGACCTACACCGAACTGAGATACCTTACAGCGTGAGCTATGAGAAAAGCGCACCGCTCC 4256
QY 5178 cgaaggagaaaggcgcagaggtatccggtaaaggcgcagggctcggaacagagagagcgac 5237
Db 4255 CGAAGGAGAAAGCGGACAGGATTCGGTAAGCGGAGGGTCGGAACAGAGAGAGCGCAC 4196
QY 5238 gaggagcttccagggggaacgcctggtatctttaaagtcctcgtcgggttccgcccct 5297
Db 4195 GAGGAGCTTCCAGGGGGAACGCCGTGATCTTTATAGTCTCTGTCGGGTTTCGCCACCT 4136
QY 5298 ctgactgagcgtcgatcttcttgatcgtcagggggcgagcgctatggaacaaacgc 5357
Db 4135 CTGACTTGAAGCGTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGAAAAACGC 4076
QY 5358 cagaacgcggcctttttacggttccgttcttggccttttggccttttgcacatgttctt 5417
Db 4075 CAGCAACGGCGCTTTTACGTTCTTGGCCTTTTGTGCGCTTTTGTCTCACATGTTCTT 4016
QY 5418 tccctgcttatccctgattcttctggataacgcgtattaccgcttcttgagtgagctgatac 5477
Db 4015 TCCTGCGTTATCCCTCATCTCTGGATACCGGTATTTACCGCCTTTTCAGTGTGATAC 3956
QY 5478 cgctcgcgcagcgcgaacaccagcagcagtcagtgagcgaagcgaagcgaagcgcg 5537
Db 3955 CGCTCGCGCAGCGGAACGACCGAGCGACGAGTGTAGCGAGGAGGAGGAGCG 3896
QY 5538 cccaatcagaacaccgctctcccgcgcttggcgcatcattcaatgaagcgtggcagca 5597
Db 3895 CCCAATACGAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAATGACAGTGGCACGA 3836
QY 5598 caggtttcccgactggaaagcgggcagtcgagcgcaacgcaattaaatgtgagttagctcac 5657
Db 3835 CAGGTTTCCCGACTGGAAAGCGGCGACGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC 3776

Qy 5658 tcattaggcaccaccagctttacacatttacttccgctcgtatgtgtgtggaattgt 5717
Db 3775 TCATTAGGACACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAAATTGT 3716
Qy 5718 gagcggaatacaatttcacacaggaacacagctatgaccatgattacgcga 5767
Db 3715 GAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 3666

RESULT 7

US-08-464-700-54
; Sequence 54, Application US/08464700
; Patent No. 6232458
; GENERAL INFORMATION:
; APPLICANT: WEISS, ANTHONY S
; APPLICANT: MARTIN, STEPHEN L
; TITLE OF INVENTION: SYNTHETIC POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,700
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL6520
; FILING DATE: 22-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PL9661
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU93/00655
; FILING DATE: 16-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GHC3USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
US-08-464-700-54

Query Match 45.6%; Score 2628.2; DB 4; Length 4045;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 2630; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3135 ttccactggccgtcttttacaacgctcgactgggaaacccctggcgttaccacacttaa 3194
Db 1 TTCACGTGCCGCTGTTTTACACGCTGCTGGAGGAAACCCCTGGCTTACCACACTTAA 60
Qy 3195 tcgccttgacacacatccctcttcgcagctggcgtaataagcagagggcccgaccga 3254
Db 61 TCGCCTTCGACGACATCCCTCTTCGCCAGCTGGCGGTAAATAGCGAGAGGCCGCCGCCGA 120

Qy 3255 tcgcccctcccaaacagttgcgcagcctgaatggcgaaatggcgccctgatgcggtattttct 3314
Db 121 TCGCCCTTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATCGGCTATTTTCT 180
Qy 3315 ccttacgcctctgtcgtattttacacacgcgcataggtgacactcagtcacaactctgc 3374
Db 181 CCTTACGCATCTGTGCGGTATTTTACACCGCATATGGTGCACTCTCAGTAGTACANTCTGCTC 240
Qy 3375 tgatccgcgcatagtttaagccagcccgacaccccgcccaacaccccgctgcgcgcctgaag 3434
Db 241 TGATGCCGCATAGTTAAGCCAGCCCGACACCCGCCCAACACCCGCTGACGCCCTTGACG 300
Qy 3435 ggcttgctcgtcccggaatccgctttacagacaagctgtgacgctctccggagcgtgat 3494
Db 301 GGCCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGAGCTGCAT 360
Qy 3495 gtgtcagaggttttcacgcctcatcccgaaacgcgcagacgaagggcctcgtatcacg 3554
Db 361 GTGTCAGAGGTTTTTACCGCTCATACCCGNAACGCGGAGACGAAAGGGCCTCGTGATACG 420
Qy 3555 cctattttataggttaatgtcatataatgggtttctcttagacgtcaggtggcacttt 3614
Db 421 CCTATTTTATAGGTTAATGTCTATATATATGTTTCTTAGACGTCAGGTGGCACTTT 480
Qy 3615 tcggggaaatgbcgcggaacccctatttgctttatttttcttaaaatataatcaaatatgca 3674
Db 481 TCGGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAATACATTCAAAATATGTA 540
Qy 3675 tcgcctcatgacacaataaccctgataatgctcaataattgaaagaagaagat 3734
Db 541 TCCGCTCATGAGACAATAACCCCTGATAATGCTTCAATAATATTGAAAAGGAAGAGTAT 600
Qy 3735 gagtattcaacatttcgctgctgcctctattcccttttttggcgcatatttgcctcctgt 3794
Db 601 GAGTATTCAACATTTCCGCTGTGCGCCCTTATTCCTTTTTTGGCGCATTTTGGCTTCTGT 660
Qy 3795 ttttgcctcaccagaaacgctggtgaaagttaaagatgctgaagatcagttgggtgcag 3854
Db 661 TTTTGTCTCACCAGAAACGCTGCTGTAAGTAAAGATGCTGAAGATCAGTTGGGTGCAG 720
Qy 3855 agtgggttacatcgaaactggatctcaacagcggtgaagatcccttgagagtttcgcgcgca 3914
Db 721 ACTGGTTACATCGAACTGGATCTCAACAGCGGTAGATCCTTGAGAGTTTTTCGCCCCCA 780
Qy 3915 agaacgttttccaatgatgagcacttttaaagttctctgctatgtggcggttattatcccg 3974
Db 781 AGAACGTTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCGGTATATATCCG 840
Qy 3975 tattgaccccgggcaagagcaactcggctgcgcgcatacactattctcagaatgaactgg 4034
Db 841 TATTGACCGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACATTGGT 900
Qy 4035 tgagtactcaccagtcacagaaaagctctacgagtgcatgacagtaagagaattag 4094
Db 901 TGAGTACTCACCAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAGAAGAAATTAG 960
Qy 4095 cagtgtgcctataacccatgagtataaacactgcggccaacttactctgacacagcgtcg 4154
Db 961 CAGTGTGCCATATACCATGAGTGATAACACTGGGCCAACTTACTTCTGACACAGGATCG 1020
Qy 4155 aggacccgaaggagctaaacgccttttttgcaacaatcgggggatcagtgaactgccttga 4214
Db 1021 AGGACCGAAGGAGCTAAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGA 1080
Qy 4215 tcgttgggaacggagctgaatgaagcattaccacacagcagcgtgacacacagtcgc 4274
Db 1081 TCGTTGGGAACCGAGCTGAATGAAGCCCATACAAACGACGAGCGTGACACACGATGCC 1140
Qy 4275 tgtagcaatggcaaacagcttgcgaactattaaactggcgcaactactactctagcttc 4334
Db 1141 TGTAGCAATGGCAACACGCTTGGGCNAACATTAATGCGGGAACACTTACTTACTCTAGCTTC 1200

|||||
Db 4435 ATGCGCATAGTTAAGCCAGCCCCGACACACCGCCCAACACCGCTGACGCGCCTGACGGG 4376
Qy 3437 ctgtctgtctcccgcatccgcttacagacaagctgtgacgctctcccgaggctcatgt 3496
Db 4375 CTGTGCTCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTCATGT 4316
Qy 3497 gtcagaggctttccacgctcaaccagaaacgcgcgagacgaaagggcctcgtgatacgc 3556
Db 4315 GTCAGAGGTTTTACCGCTCATCACGAACCGCGAGACGAAGAGGCGCTCGTGATACGCC 4256
Qy 3557 tatctttataggtaaatgctatgataaataaggtttctctagacgctcaggtgacatttc 3616
Db 4255 TATTTTTATAGTTAANTGTCATGATAAATAGTGTTCCTTAGACGTACAGTGGCACTTTTC 4196
Qy 3617 ggggaaaagtgcgggaaacccctatttgtttatttttcaataacatcaaatatgtatc 3676
Db 4195 GGGGAAATGTGCGCGAAACCCCTATTGTGTTATTTTTCTTAATACATCAAAATATGTATC 4136
Qy 3677 cgtcatgagacaataaacctcgataaaagcttcaataatattgaaaaagggaagataga 3736
Db 4135 CGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATTTGAAAAAGGAGATATGA 4076
Qy 3737 gtattcaaatctccgtgcgccttattcccttttttgcggcattttgcctctcgttt 3796
Db 4075 GTATTTCAACATTTCCGTGTGCGCCCTTATTCCTTTTTTTCGGCATTTTTGCCCTCTCTGTTT 4016
Qy 3797 ttgtcccccagaaacgctggtgaaagtaaaagatagctgaagatcagttgggtgcacgag 3856
Db 4015 TTGCTTCACCCAGAAACCGTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 3956
Qy 3857 tgggttacatgaaactgatctcaacagcggtaagatccttgagagtttccgcccgaag 3916
Db 3955 TGGGTATACATCGAATCGATCTCAACACGCGTAAGATCCTTGAGAGTTTTTCGCCCGGAAG 3896
Qy 3917 aacgttttccaatgatgagcacttttaagttcctgctatgltggcgcgtattatcccgta 3976
Db 3895 AACGTTTTCCAATGATGAGCACTTTTTAAAGTTCGCTATGTGGCGCGGTATTTATCCCGTA 3836
Qy 3977 ttgacgcggggaagagcaactcgtgcgccgatacaactattctcagaatgacttgggtg 4036
Db 3835 TTGACGCGGGGAGAGCAACTCGGTTCGCCGATACACTATTCTCAGAATGACTTGGTTG 3776
Qy 4037 agtactccagctcacagaaaagaactcttaaggatgagatgagcagatgagagaattatgca 4096
Db 3775 AGTACTACCACTACAGAAAAGCATCTACGGATGGGATGAGATGAGATAAGAGAAATTATGCA 3716
Qy 4097 gtgctgccataaccatgagtataaactcgtgcgcaacttactctgacaacgacgagag 4156
Db 3715 GTGCTGCCATTAACCATGAGTATACACTGCGGCCAACTTACTTCTGACAACGATCGGAG 3656
Qy 4157 gaccgaaggagctaaccccttttttgcacaacatgggggataatgtaactcgccttgatc 4216
Db 3655 GACCCGAAGAGCTAAACCGCTTTTTTGCAACAATGGGGGATCATGTAACCTCGCCCTTGATC 3596
Qy 4217 gttgggaaccggagctgaatgaagacataaccaaacgagcagcgtgacaccagatgcctg 4276
Db 3595 GTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCCAGATGCCCTG 3536
Qy 4277 tagcaatggcaacaacglttgcgaacactattaaactggcgaaacttactctagcttccc 4336
Db 3535 TAGCAATGGCAACAACGTTGCCCAACTATTAACTTGGCGAACTACTTACTTAGCTTCCC 3476
Qy 4337 ggcacaataatagactggatggaggcggaataaagtgcaggaccacttctgcctcgg 4396
Db 3475 GGCAACAATTAATAGACTGGATGGAGGGGATAAAGTTGCAGGACCACTTCTGCGCTCGG 3416
Qy 4397 cccctccggctgctgggtttatgtctgataaaatctggagcgggtgagcgtgggtctcgg 4456
Db 3415 CCCTTCCGGCTGGCTGGTTATTTGTGTGATAAATCTGGAGCCGCTGAGCGTGGGTCTCGGG 3356
Qy 4457 gtatcttgcagcactgggggcagatggtgaagccctcccgatcgtattatctacaga 4516
|||||

Db 3355 GTATCATTTGCACACACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 3296
Qy 4517 cggggagtcaggaaactatggatgaacgaaaatagacagatcgtctgagataggtgcctcac 4576
Db 3295 CGGGGAGTCAGGCAACTATGGATGAAGAAATAGACAGATCGCTGAGATAGTGCTCAC 3236
Qy 4577 tgatataagcattgggtaactgctcagaccgaagttaactcatatatactattagattgataa 4636
Db 3235 TGATTAAGCATTTGGTAACGTGTGAGCAACCAAGTTTACTCATATATACTTTAGATTGATTTAA 3176
Qy 4637 aacttcatttttaattaaaagactaggtgaagatccctttttgataatctcctcagacca 4696
Db 3175 AACTTCATTTTTTAATTTAAAGATCTAGTGAAGATCCTTTTTGATAATCTCATGACCA 3116
Qy 4697 aaatcccttaacgtaggttttccactgagcgtcagaccccgtagaagaagatacaag 4756
Db 3115 AAATCCCTTAACGTGAGTTTTCCTTCCACTGAGCGGTGAGACCCCGTAGAAAAGATCAAAAG 3056
Qy 4757 gatctcttgagatcccttttttctgcgagtaactgctgctgttgcaaaacaaacacacac 4816
Db 3055 GATCTTCTTGAGATCCTTTTTTCTGCGGTAACTGCTGCTTGCMAACAAAAAACCCAC 2996
Qy 4817 cgtaccagcgggtggtttgtttgcgggatacaagagctaccaactcttttccgaaagtaa 4876
Db 2995 CGCTACCAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCCGAAAGTAA 2936
Qy 4877 ctggcttcagcagcagcagataaccaatactgctctcttagttagccgttagtaggc 4936
Db 2935 CTGGCTTCAGCAGAGCGCAGATACCAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCC 2876
Qy 4937 accaactcaagaactctgtagcaccgcctataactcgtctgtctgtctgtctgtctgtctgt 4996
Db 2875 ACCACTTCAAGAACTCTGTAGCACCCTTACATACCTCGCTCTGCTTAATCTCTTACCAG 2816
Qy 4997 tggctgtccagtgagcagataagtcgtcttaccgggttgactcaagacagatagttac 5056
Db 2815 TGCTGTCTGCCAGTGGCGATAGTCTGTCTTACC GGTTGGACTCAAGACGATAGTTAC 2756
Qy 5057 cgggtaagggcagcggctcgggctgaacgggggggttcgtgcacacagcccgcttgagac 5116
Db 2755 CGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCGTGCACACAGCCAGCTTGGAGC 2696
Qy 5117 gaaacactaacccgaactgagataacctacacgctgagctatgagaagcgcacgcttc 5176
Db 2695 GAACACCTTACACCACTGAGATACCTACAGCTGAGCTATGAGAAAGCGCACGCTTC 2636
Qy 5177 ccgaaggagaaaggcggacaggtatccggttaagcggaggtcggaaacagagagcgcga 5236
Db 2635 CCGAAGGAGAAAGCGGACAGGTATCCGTAAGCGGAGGGTTCGGACAGAGACCGCA 2576
Qy 5237 cagaggagcttccaggggaaacgcctggtatctttatagtcctgctcggttttcgccaac 5296
Db 2575 CGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGTCTGTCGGGTTTCGCCACC 2516
Qy 5297 tctgactgagcgtcgatttttctgctgagcgtgagggggggcgagccttatggaaaaag 5356
Db 2515 TCTGACTTGAGCGTCGATTTTGTGATGCTGCTCAGGGGGGCGAGCCCTATGGAAAAACG 2456
Qy 5357 ccagaaacgcggcctttttacggttccctggccttttctgctggtcttttgcacatgtct 5416
Db 2455 CCAGCAACCGCGCCTTTTTTACGGTTCTCGCCCTTTTGTCTGGCCTTTTGTCTACATGTTCT 2396
Qy 5417 ttctcgtgcttaccctgattctgttgataaacgcttattaccgcttttgagtgagctgata 5476
Db 2395 TTCTCTGCTTATCCCTGATCTGTGATAACCGTATTACC GCTTTTGAGTGAGCTGATA 2336
Qy 5477 ccgctcgcgcagcagcagaaacgagcagcagcagcagcagcagcagcagcagcagcagc 5536
Db 2335 CGCTCGCCGCGACGCGAACGACGCGAGCGAGTCACTGAGCGAGGAGGAGCGGAAGAGC 2276
Qy 5537 gcccaatagcaaacgcgcctctcccgcgcgttggccgcttcatattaatcagctgacag 5596
Db 2275 GCCCAATACGCAAAACCGCCTCTCCCGCGCGGTTGGCCGATTCATTAATGACGTGCACG 2216
|||||

Qy	5597	acaggtttcccgactggaagcggcgagtgagcgcaacgcgcaataaagtgtgattgagctca	5656
Db	2215	ACAGGTTTCCCGACTGAAAGCGGGCAGTGACGCGCAACGCAATTAATGTGAGTTAGCTCA	2156
Qy	5557	ctcattaggcaaccccgaggtttacacotttatgcttcctcgctcgctgatgttgtgtggaattg	5716
Db	2155	CTCATTAGGCACCCAGGCTTTACATTTATGCTTCCGGCTCGTATGTTGTGGGAATTG	2096
Qy	5717	tgaacggataacaatttcacacagaaacagctatgaccatgattaacg	5764
Db	2095	TGAGCGGATTAACAATTTTCACACAGAAACAGCTATGACCATGATTACG	2048

RESULTS

```

US-09-147-208-26/c
: Sequence 26, Application US/09147208
: Patent No. 6333303
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Antiviral Ricin-Like Proteins
: NUMBER OF SEQUENCES: 71
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BERESKIN & PARR
: STREET: 40 King Street West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5H 3Y2
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/147,208
: FILING DATE: 02-MAR-1999
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Rudolph, John R.
: REGISTRATION NUMBER: 38,003
: REFERENCE/DOCKET NUMBER: 7841-76
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 364-7311
: TELEFAX: (416) 361-1398
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9639 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-09-147-208-26

```

```
Query Match      45.68; Score 2628; DB 4; Length 9639;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	3137	cactggccgctcgcttttacaac	ctgactgactgggaaacccctggcgttaccaccaacttaac	3196
Db	9637	CACTGGCGTCGTTTTACAAC	TCGTGACTGGGAAACCCCTGGCGTTACCCAACTTAATC	9578
Qy	3197	gccttgagcacatccccctt	tcgcagctggcgtaatagcgaagagggccgcacgcgac	3256
Db	9577	GCCTTGCGACACATCCCCCT	TTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCACCGATC	9518
Qy	3257	gccttcccaacagttgcgcag	cctgaatggcgaaatggcgctgatcggtattttctcc	3316
Db	9517	GCCTTCCCAACAGTTGCGC	AGCCTGAATGGCGAATGGCGCTGATGGGTATTTTCTCC	9458
Qy	3317	ttacgcaatctgtcggtattt	ccacacgcgcaatatggtcactctcagtcacaactctgctctg	3376

Db	9457	TTACGCATCTGTGCGGATATTTTACACCGCATATGGTGCACATCTCTACGTACAAATCTGCTCTG	9399
Qy	3377	atgcgcatagtttaagccagcccgacacccgcaacaccccgctgacgcgcctcgacggg	3436
Db	9397	ATGCCGATAGTTAAGCCAGCCCGACACCCGCCACACCCCGTGTACGCCCTTGACGGG	9338
Qy	3437	cttgctgtcccgggcatccgcttacagacaaagtctgacccgtctccggagagctgcattg	3496
Db	9337	CTTGCTGTCTCGCGCATCCGCTTACAGACAAGCTGTGACCGTCTCGGAGCTGCATGT	9278
Qy	3497	gtcacagggttttcaacogtcatcacccgaacgcgcgcagacgaagggcctcgtagacgcc	3556
Db	9277	GTACAGAGTTTTACCGTCTATCACCGAAACGCGCAGACGAAAGGCGCTCGTGATACGCC	9218
Qy	3557	tattttataggttaagtcatgataataagtgtttcttagacgtcaagtgggcacatttc	3616
Db	9217	TATTTTATAGGTTAATGTCATGATAATATGTTCTTTAGACGTCAAGTGGCATTTTC	9158
Qy	3617	ggggaaatgtgcgggaacccctattgtttattttcttaaatcacattcaaatgtatc	3676
Db	9157	GGGGAATGTCGCGAACCCCTATTGTGTTATTTTTCATAATTCATAATGTATC	9098
Qy	3677	cgctcatgacacaataaacccctgataaagtctcaataatatgaaaaaggaagtatga	3736
Db	9097	CGCTATGAGACAAATAACCTCTATAAATGCTTCAATAATATTGAAAGGAAGATGA	9038
Qy	3737	gtattcaacatttcggtgcgccttatcccttttttgcggcattttgccttcctgttt	3796
Db	9037	GTATTCAACATTTCCGTGTGCGCCTTATTCCCTTTTTCGGGCATTTTGCTTCCGT	8978
Qy	3797	ttgctcacccagaaacgcgtggtagaagtaaaagatgctgaagatcagttgggtgcacgag	3856
Db	8977	TTGCTCACCCAGAACGCTGGTGAANGTAAAGATGCTGAAGATCAGTTGGGTGCAGGAG	8918
Qy	3857	tgggttacatcaactggatctcaacacagcggtaagatccttgagagtbtctgcgccgaag	3916
Db	8917	TGGGTATACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTTCGCCCGC	8858
Qy	3917	aagcttttccaatgatgagcacttttaagttctgctatgtggcgcggatattatccgta	3976
Db	8857	AACGTTTCCATATGATGAGACATTTTAAAGTTCTGCTATGTGGCGGGTATTATCCGTA	8798
Qy	3977	ttgacgcgggcaagagcaactcgtgcgcgcatacactattcttcagaaatgacttggttg	4036
Db	8797	TTGACGCGCGCAAGAGAACCTCGGTGCGCGCATACATATTCTCAGATGACTTGGTTG	8738
Qy	4037	agtactcaccgttcacagaaaaagcatcttacggatggcatgacagtaagagaattatgca	4096
Db	8737	AGTACTCACCGTCAACAGAAAAGCATTTACGGATGGCATGACAGTAAGAGAAATTATGCA	8678
Qy	4097	gtctgcctaaccatagtgataacacitcgggccaacttactctgacaacgatcggag	4156
Db	8677	GTGCTGCCATACCATGATGATGAACACATGCGGCCAACATTACTTCTGCAACGATCGGAG	8618
Qy	4157	gaccgaaggagctaacccgttttttgcacaacaatggggatcatgtaaactcgcccttgatc	4216
Db	8617	GACCGAAGGAGCTAACCGCTTTTTCACACACATGGGGATCATGTAACTCGCCTTGATC	8558
Qy	4217	gttggaaacccgagctgaatgaagccataccaaacacacagcagcgtgacacaacgatgcctg	4276
Db	8557	GTTGGAAACCGGAGCTGAATGAAGCCCATACCAACACACGAGCGCTGACACACGATGCCCTG	8498
Qy	4277	tagcaatggcaaacgcttgcgaactaactgagcgaactacttacttacttagctccc	4336
Db	8497	TAGCAATGGCAACAGCTTGGCGGAACCTATTAACTGGCGAACATTACTCTAGCTTCCC	8438
Qy	4337	ggcaacaataatagactggatggagcgggataaagtgtgacgacacacttctcgctcgg	4396
Db	8437	GGCAACAATTAATAGACTTGSATGGAGCGGATAAACTTCGAGGACCACTTCTCGGCTCGG	8378
Qy	4397	cccttcggcgtgcgtgttttatgtcgtataaatactctggagcccggtgagctctcgog	4456
Db	8377	CCCTTCCGCGCTGGCTGGTTTATTGCTGATTAATCTCGAGCGGTGAGCTGGGTCTCGCG	8318

Qy 4457 gatacattgcaagcaactggggccagagatggtaagccctcccgatcgtagttatctacaga 4516
Db 8317 GTATCATTTGACGACTGGGGCCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 8258
Qy 4517 cggggagtcaggcaactatgataaacgaaatagacagatcgctgagataggtgctcac 4576
Db 8257 CGGGAGTCAGGCAACTATGATGNACGAATAGACAGATCGCTGAGATAGTGCTCAC 8198
Qy 4577 tgaataagcattggtgaactgtcagaccagtttactctatatactatttagattttaa 4636
Db 8197 TGATTAGCATTTGGTAACGTGACAGCAAGTTTACTCATATATATCTTTAGATTGATTAA 8138
Qy 4637 aacttcatttttaatttaaaagatctaggigaagatcccttttttgataatctcatgaca 4696
Db 8137 AACTTTCAATTTTAAATTAATAAGAGTCTAGGTGAAGATCCCTTTTGTATATCTCATGACCA 8078
Qy 4697 aaatcccttaacgtgagtttttcctcactgagcgctcagaccctgtagaagaatcaag 4756
Db 8077 AAATCCCTTAACGTGAGTTTTCGTTCCTACTGAGCGTCAGACCCCGTAGAAAAGATCAAG 8018
Qy 4757 gatctcttgagatcccttttttttgcgcgttaactctgctgtgcacaaacaaacccac 4816
Db 8017 GATCTCTTGAGATCCTTTTTCGCGCTAATCTGCTGTGCAACAAAAAACCCAC 7958
Qy 4817 cgctaccagcggttggtttgttgcggatcaagagctaccactcttttcggaaggtaa 4876
Db 7957 CGCTACCAGCGGTGTTTGTTCGGGATCAAGAGCTACCAACTCTTTTTCGGAAGGTAA 7898
Qy 4877 ctggtctcagcagagcgatatacaactgttcctctagtgtagcgttagtgagcc 4936
Db 7897 CTGGCTTCAGCAGAGCGGATACCAATAGTCTCTCTAGTGAGCCGTAGTTAGGCC 7838
Qy 4937 accactcaagaactctgttagcagcgctacatacctcgctcgtgtaactcctgtttaccag 4996
Db 7837 ACCACTTCAAGAAGTCTGTAGCAGCGCTACATACCTCGCTCTGCTAATCTGTTTACCAG 7778
Qy 4997 tggctctgcagtgaggatagtcgtctctaccgggttggactcaagagatagttac 5056
Db 7777 TGGCTGCTGCCAGTGGCGATGAAGTCGTCTTACCGGGTGGACTCAAGACGATAGTTAC 7718
Qy 5057 cggataagcgccagcgggctgggctgaacggggggttcgtgcacagcccagcttgagac 5116
Db 7717 CGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTCGTGCACACAGCCAGCTTGGAGC 7658
Qy 5117 gaacgactacacggaactgagatgagatgagcgtgagctatgagaaagccagcgttc 5176
Db 7657 GAACGACCTACACGGAAGTACGATACCTACAGCGTGAGCTATGAGAAAGCGCCAGCTTC 7598
Qy 5177 ccgaaggagaaagcgagacaggtatccggtaagcggcagggctcggaacagagagagcga 5236
Db 7597 CCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGGTCCGAAACAGGAGAGCGCA 7538
Qy 5237 cgaggagcttccagggggaaacgctgtgtatctttagtctcgtcggttctgcgccacc 5296
Db 7537 CGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTCGGTTCGCCAC 7478
Qy 5297 tctgactgagcgtgatattttgtgactgctcagggggcgagcctatggaaaaagc 5356
Db 7477 TCTGACTTGAGCGTCGATTTTGTGTATGCTGTCAGGGGGCGGAGCCTATGGAATAACG 7418
Qy 5357 ccagcaacggcgcttttttaccggttctcgcccttttgcgtgaccttttgcacatgttct 5416
Db 7417 CCAGCAACGGCGCTTTTACGGTTCCTGGCCCTTTTGTGCTGCTTCTCACATGTTCT 7358
Qy 5417 tctctgcttatccctgattctgtggataaccgtattaccgcttttagtgagctgata 5476
Db 7357 TTCTGCGCTTATCCCTGATTCTGTGGATACCGTATTTACCGCCTTTGAGTGAGCTGATA 7298
Qy 5477 ccgctcgccgagccgaacccagcagcgagtcagtgagcaggaagcgaagagc 5536
Db 7297 CCGCTCGCCGAGCGGAACGACCGAGCGACGAGTCAGTGAGCGAGGAGCGGAAGAGC 7238

Qy 5537 gccaaatcgcaaaacccgctctcccgcgcttgccgattcaatcaatcgagctggcaag 5596
Db 7237 GCCCAATACGAAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGACAGCTGCACG 7178
Qy 5597 acaggtttcccgactggaagcgagcgagtcgagcgcaagcaatcaatgtgagttgactca 5656
Db 7177 ACAGGTTTCCGACTTGAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7118
Qy 5657 ctcatgagccacccagcgttttacactttatgctctccggtcgatgtgtgtggaattg 5716
Db 7117 CTCATTAGGACCCCGAGGCTTTACACTTTATGCTTCCGCTCGTATGTTGTGTGGAATTG 7058
Qy 5717 tgagcgataaacaatttcacacaggaacacagctatgacacgattacag 5764
Db 7057 TGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACG 7010

RESULT 10
US-08-232-016-23/c
; Sequence 23, Application US/082332016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELISSEN, Marc
; APPLICANT: SOETAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN AARSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pPS0212
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1785
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated modified bt2 (cryIAb) gene, also designated as t

```
OTHER INFORMATION: cryIab6 gene."
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1793..2026
  OTHER INFORMATION: /note= '3' regulatory sequence
  OTHER INFORMATION: containing the polyadenylation site derived from Agrobacteriu
  OTHER INFORMATION: T-DNA gene 7."
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 2396..2921
  OTHER INFORMATION: /note= '35S promoter sequence
  OTHER INFORMATION: derived from Cauliflower mosaic virus."
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 2922..3581
  OTHER INFORMATION: /note= "Coding sequence of
  OTHER INFORMATION: chloramphenicol acetyl transferase gene."
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 3582..4407
  OTHER INFORMATION: /note= '3' regulatory sequence
  OTHER INFORMATION: containing the polyadenylation site derived from Agrobacteriu
  OTHER INFORMATION: T-DNA octopine synthase gene."
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 5600..6457
  OTHER INFORMATION: /note= "Sequence complementary to
  Patent No. 5952547
  OTHER INFORMATION: the coding sequence of the beta-lactamase gene."
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 7071..7566
  OTHER INFORMATION: /note= "Trl' and Tr2 promoter
  OTHER INFORMATION: derived from Agrobacterium T-DNA (with modified leader with
  OTHER INFORMATION: respect to sequence of pJD884 of SEQ ID NO. 22."
US-08-232-016-23

Query Match 45.6%; Score 2627; DB 2; Length 7566;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3124 ctgaagcgaattcactggccgtcggtttacacacgctgactgggaaacccctggcgtt 3183
|||
DB 7066 CTGCACCCGAAGCTTACTGGCCGTCGTTTACACACGTCGTGACGCGGAAACCCCTGGCGTT 7007

QY 3184 acccaacttaatgccttcgacacacatccccctttcgccagctggcgtaatagcgaagag 3243
|||
DB 7006 ACCCAACTTAATGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAG 6947

QY 3244 gccgcacccgatgcgcttcccaaacagttgcgagcgtgaatggcgccctgatg 3303
|||
DB 6946 GCCCGCACCGATGCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGCCCTGATG 6887

QY 3304 cggattttctctacgcacatctgctggtatttcacacgcgatattggtgcactctcagt 3363
|||
DB 6886 CGGTATTTCTCTTACGATCTGTGCGGTATTTACACCGCATATGTTGCACTCTCAGT 6827

QY 3364 acaatctgctgatccgcgatagtttaagccagccccgcacacccgcacaccccgctgac 3423
|||
DB 6826 ACAATCTGCTCTGATCGCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCCGCTGAC 6767

QY 3424 ggcgcctgacggcctgtctgtcccgcatccgcttacagacaagctgtgacgcgtctcc 3483
|||
DB 6766 GCGCCCTGACGGGCTTGTCTGCTCCCGGATCGGTTTACAGCAAGCTGTGACCGTCTCC 6707

QY 3484 gggagctgcatgtgtcagaggttttccacgctcatccacccgacccgagacgaagggc 3543
|||
DB 6706 GGGAGCTGCATGTGTGAGAGTTTTCACCGTCATCACCCGAAACCCGCGAGACGAAGGGC 6647

QY 3544 ctcgtagacgcctattttataggttaattgacataataatggtttcttagacgtca 3603
|||
DB 6646 CTGCTGATAGCGCTATTTTATAGGTTAATGTGTCATGATAATAATGTTCTTTAGACGTCA 6587
```

```
QY 3604 ggtggcaacttttcggggaaaatgtgcgcggaacccccctatttggtttttctataaacat 3663
|||
DB 6586 GGTGGCACTTTTCGGGAAAATGTGCGGAAACCCCTATTGTTGTTTATTTTCTAAATACAT 6527

QY 3664 tcaaatatgatccgtcatgcagacacatccctgtataaatgcttcaataattgaaaa 3723
|||
DB 6526 TCAAAATATGATCCGCTCATGAGACAAATAACCCCTGATAAATGCTTCAATAATATTGAAA 6467

QY 3724 aggaagatgatgatttcaacatttccgtgtcgcccttattcccttttttggcgcat 3783
|||
DB 6466 AGGAAGATGATGATTTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTGGCGCATTT 6407

QY 3784 tgccttccgttttttgcctaccacgaacgctggtgaaagtataaagtgcgtgaagatcag 3843
|||
DB 6406 TGCCTTCCGTGTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAG 6347

QY 3844 ttgggtgcacgagtggtttacatcgaactggaatcctcaacagcgtgaagatcccttgag 3903
|||
DB 6346 TTGGGTGCACGAGTGGTTACATCGAATCGAATGATGAGCAGTCTTCAACAGCGGTAAAGAT 6287

QY 3904 ttctgccccgaagaacgcttttccaatgatgagcactttttaaagttctgtctatgtggcg 3963
|||
DB 6286 TTTGCGCCCGGAAGACGTTTTCCAATGATGAGCAGTCTTAAAGCTTCTGCTATGTGGCGG 6227

QY 3964 gtattatccgctattgacgcggcggaagcgaactcggtcgcgcgacatacactattctcag 4023
|||
DB 6226 GTATTATCCGTTATTGACGCGCGGGAAGAGCAACTCGGTGCGCGCATACACTACTTCTCAG 6167

QY 4024 aatgacttggttgactcaccagtcacagaaaagacatcttcacgatggcagcagta 4083
|||
DB 6166 AATGACTTGGTTGAGTACTCACCAGTACACAGAAAGATCTTACGGATGGCATGACAGTA 6107

QY 4084 agagaattatgcgtgctgcataaacatgagtataaacactgcgcgaacttactcttg 4143
|||
DB 6106 AGAGAATTATGCAGTCTGCCATAACCATGAGTGATAAACACTGCGGCCAATCTTACTCTG 6047

QY 4144 acaacgatcggagaccgaaggagctaacgcgtttttgcacaacatgggggataatgta 4203
|||
DB 6046 ACAACGATCGGAGACCGAAGGAGCTTAACCGCTTTTTCACAACTATTTGCGGAGTATGTA 5987

QY 4204 actcgccttgatcgttgggaacccgagcgtgaatgaagcaccataccaaacgacgagctgac 4263
|||
DB 5986 ACTCGCTTGATGCTTGGGNAACCGGAGCTGTAATGAAGCCATACCAACGACGAGCGTGAC 5927

QY 4264 accacgatgcctgtagcaatggcaaacagctgtgcgaactattaaactggcgaactactt 4323
|||
DB 5926 ACCACGATGCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTACTT 5867

QY 4324 actctagcttcccggaacaaatataatagactggatggagcgggataaaagtgcaggacca 4383
|||
DB 5866 ACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCGAGACCA 5807

QY 4384 ctctcgcctcggcccttcgcgctggtttatttgctgataaaactctggagcggtag 4443
|||
DB 5806 CTCTCGCTCGCCCTTCGCGCTGGCTGGTTTATGCTGTAATAATCTGGAGCGCGGTGAG 5747

QY 4444 cgtgggtctcgggtatcatcttcgacgacctggggccagatggtgaagccctccctatcgta 4503
|||
DB 5746 CGTGGGTCTCGCGGTATCATTTGAGCAGCTGGGGCCAGATGGTAAGCCCTCCCGTATCGTA 5687

QY 4504 gttatcacgacgggggagtcaggaactatggaacgaataatagacagatcgcgtgag 4563
|||
DB 5686 GTTATCTACACGACGGGGAGTCAGGCAACTATGGTAAGCAAAATAGACAGATCGCTGAG 5627

QY 4564 ataggtgcctcactgattgaagcattgtaactctcagaccacttactactatatactt 4623
|||
DB 5626 ATAGGTGCCTCACTGATTAAGCATTGGTAATCTGTCAGACCAAGTTTACTCATATATACTT 5567

QY 4624 tagattgattaaaaactcatttttaattaaagatcctagtggaagatcccttttgat 4683
|||
DB 5566 TAGATTGATTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTAT 5507
```

QY 4684 aatctcatgaccacaaatcccttaacgtgagttttcttccactgagcgtcagaccccgta 4743
Db 5506 AATCTCATGACCAAAATCCCTTAACGTGAGTTTTCCTCACTGAGCGTCAGACCCCGTA 5447
QY 4744 gaaaagatacaaggatctcttgagatcccttttttctgcggttaactgtgctgcttgca 4803
Db 5446 GAAAAGATCAAAAGGATCTCTTTGAGATCCTTTTTTCTGCGGTAAATCTGCTGTGCAA 5387
QY 4804 acaaaaaaacacccctaccagcgttggttttgcgagataagagagctaccaactctt 4863
Db 5386 ACAAAAAAACCCACCTACCACGCTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTT 5327
QY 4864 ttccgaaggtaactggtctcagcagcagataccacaaatctcttctctagttag 4923
Db 5326 TTTCCGAGGTAACTGGCTTCACGAGCGCAGATACCAAAATCTCTCTTAGTGTAG 5267
QY 4924 ccgtagttaggccacacttcaagaactctgtagcacccctacatacctcgtctgcta 4983
Db 5266 CCGTAGTTAGGCCACACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTA 5207
QY 4984 atcctgttaccagtgtgctgcagtgagtgataagtgctgtcttaccgggttgactca 5043
Db 5206 ATCCCTGTTACCAAGTGGCTCTCCACAGTGGCGATAAGTCGTCTTTACCGGTTGGACTCA 5147
QY 5044 agacgatagttaccgataaagcgcagcgtcggtcgactaaacgggggttcgtgcacacag 5103
Db 5146 AGACATAGTTACCGATTAAGCGCAGCGTCCGGCTGAACCGGGGTTTCGTGCACACAG 5087
QY 5104 cccagcttggagcgaacacacctacacccgaactgagatacctcacagcgtgagctatgagaa 5163
Db 5086 CCCAGCTTGGAGCGAAGCAGCTACACCGAACTGAGATACCTACACGCTGAGCATTGAGAA 5027
QY 5164 agcgcacagcttccgaaggaggaagcggacaggtatcccggttaagcggggtcgga 5223
Db 5026 ACCGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGGTGCGA 4967
QY 5224 acagagagcgcagcagagagcttccaggggaaacgcctgatatctttatagctctgc 5283
Db 4966 ACAGAGAGCGCAGAGGAGGCTTCAGGGGGAACCGCTGTGTATCTTTATAGTCTGTGC 4907
QY 5284 ggggttcgcaactctgactgagcgtcgattttgtgatcgtcgtcagggggcgagc 5343
Db 4906 GGGTTTGGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGC 4847
QY 5344 ctatggaaaaacgcagaaacgcgcttttcaagtttctggttcttgcgttcgtggttttt 5403
Db 4846 CTATGGAAAAACGCCAGCAACCGCGCTTTTACGGTTCTCGCCCTTTTGTCTGGCCTTTT 4787
QY 5404 gctcacatgttcttctcgtcgttatccctgattctgtgataacggtattaccgcttt 5463
Db 4786 GCTCATATGTTCTTCTCGCTTATCCCTGTATCTGTGGATAACCGTATTACCGCCTTT 4727
QY 5464 gagtgcgtgatcacgctgcgcgacgcgaacgcagcgcagcgcagcgcagtcagtgagcgcag 5523
Db 4726 GAGTGAGCTGATACCGCTCGCGCGAGCGCAACGACGAGCGAGTCAGTGAGCGAG 4667
QY 5524 gaagcgggaagcgcgcatacgaacacgcctctcccgcgcttggtggcgatttcattaa 5583
Db 4666 GAAGCGGAAGAGCGCCCAATACGAAACCGCTCTCCCGCGCTTGGCCGATTCATTAA 4607
QY 5584 tgacgtggcagcagaggtttcccgactggaagcgggcagtgagcgaacgaactaat 5643
Db 4606 TCCAGCTGCGACAGCAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAAT 4547
QY 5644 gtgagtgcactcatttagcaccccgaggtttacattttatgcttcgcgctgctatg 5703
Db 4546 GTGAGTGTAGCTCATTATAGCACCCCGAGGCTTTACATTTTATGCTTCGGCTCGATG 4487
QY 5704 ttgtgtggaattgtgagcggatacaacatttcacagggaaacagcgtatgacctatgac 5763
Db 4486 TTGTGTGGAATTGTGAGCGGATAAACAATTTTCACAGGAAACAGCTATGACCATGATTAC 4427
QY 5764 gcc 5766

Db 4426 GCC 4424
|||
RESULT 11
US-08-232-016-22/c
; Sequence 22, Application US/08232016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELIJSSEN, Marc
; APPLICANT: SOETAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN ARSSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; TITLE OF INVENTION: PLANT CELLS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pJD884
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1869
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated bt2 (cryIab) gene, also designated as the bt884
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1877..2110
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA gene 7."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2480..3005
; OTHER INFORMATION: /note= "35S promoter sequence
; OTHER INFORMATION: derived from Cauliflower mosaic virus."
; FEATURE:
; NAME/KEY: misc_feature

```

: LOCATION: 3006...3665
: OTHER INFORMATION: /note= "Coding sequence of
: OTHER INFORMATION: chloramphenicol acetyl transferase gene."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3666...4491
: OTHER INFORMATION: /note= "3' regulatory sequence
: OTHER INFORMATION: containing the polyadenylation site deriv
: OTHER INFORMATION: T-DNA octopine synthase gene."
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 5684...6541
: OTHER INFORMATION: /note= "Sequence complementary to
: Patent No. 5952547
: OTHER INFORMATION: the coding sequence of the beta-lactamase
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 7155...7639
: OTHER INFORMATION: /note= "Trl' and TR2' promoter
: OTHER INFORMATION: derived from Agrobacterium T-DNA."
: US-08-232-016-22

```

Query Match	45.6%	Score 2627;	DB 2;	Length 7639;
Best Local Similarity	99.6%;	Prod. No. 0;		
Matches 2633;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
Qy	3124	cttgaagcgaattcaactgcccgtctgtttcaacqctcgtactgctggaaaaacccctggcgtt	3183	
Db	7150	CTGCAGCAAGCTTACGTGGCCGCTGTTTTACACGTCGTGACTGGGAAAAACCTGGCGGT	7091	
Qy	3184	accbaacttaactgccttgcgagcacatccccctttcgccagctggcgttaatagcgaag	3243	
Db	7090	ACCCAACCTAATCGCTTGCAGCACATCCCCCTTTGCGCAGTGGCGTAATAGCGAAG	7031	
Qy	3244	gccgcaccgcatcccttcccaacagttgcgcagcctggaatggcgaatggcgctgatg	3303	
Db	7030	CCCGCACCGATCGCCCTTCCCAACAGTTGCGAGCCTGAATGGGGAATGGCCCTGATG	6971	
Qy	3304	cggattttctccttaacgcatctgtgcggtatttcacaccgcatatggctcactctcagt	3363	
Db	6970	CGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACACCGCATATGTCACATCTCAGT	6911	
Qy	3364	acaatcgtctctgatgcgcagtagttaagcagcccccgaccccccgccaaaccccgctgac	3423	
Db	6910	ACAATCTGCTCTGATGCGCGCATAGTTAAGCAGACGCCCGCACACCCGCCAACACCCCGCTCAC	6851	
Qy	3424	gcgcctgaecgggctgtctgctcccgscatccgctttacagacaagctgtgacqctctcc	3483	
Db	6850	CGGCCCTGACGGGCTTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCC	6791	
Qy	3484	gggagctgcatagtctagagaggttttcacggtcactccacccgaaacgcgcgagacgaagggc	3543	
Db	6790	GGGAGCTGCATGTGTACAGAGGTTTTACCGTCTATCCACCGAAACGCCCGCAGACGAAAGGCG	6731	
Qy	3544	ctcgtgacgcctattttatagtgtaactgatgataataatggttcttttagacgtca	3603	
Db	6730	CTCGTGATACGCCCTATTTTATAGGTTAATGTCTAGTAAATAAGTGGTTCTTAGACGTCA	6671	
Qy	3604	ggtggcacttttcggggaatgtgcgcgggaacccctattgtttatttttcttaatacat	3663	
Db	6670	GGTGGCACTTTTTCGGGGAATGTGCGCGGAACCCCTATTGCTTTATTTTCTAATAACAT	6611	
Qy	3664	tcaaatatgatcgcgtcatagagacaataacccctgataaatgtcttcaataattgaaaa	3723	
Db	6610	TCAAATATGTATCGCTCATGAGACAAATAACCCCTGATAAATGCTTCAATAATATTGAAAA	6551	
Qy	3724	aggaagatgatgattcaacatttcgcgtgcgcccttatcccttttttgcgcgcat	3783	
Db	6550	AGGAAGATATGAGTATTTCAACATTTCCGTGTGCCCCATTATCCCTTTTTTTCGGCATTT	6491	
Qy	3784	tgccttcctgttttgcaccaccgaaacgcgtggtgaaagtataaagatgcgtgaagatcag	3843	

Qy	4924	ccgtagttaggcaccacacttcaagaactctgtagcacgcgctacatacctcgtctgtcta	4983
Db	5350	CCGTAGTtagGCCACACACTTCAAGAACTCTGTAGCACGCGCTACATACCTCGCTCTGCTA	5291
Qy	4984	atccttaccagtgctcgtccagtggcgataagtcgtctcaccgggttgactca	5043
Db	5290	ATCCTGTTTACCAGTGGCTGCTCCAGTGGGATAAGTCGTGCTTACCGGGTGGAGTCA	5231
Qy	5044	agacgatagttaccggataaggcgacggctcgggctgaacggggggttcgtgcacacag	5103
Db	5230	AGACGATAGTTCACGGATAAGCGCAGCGGTGGGCTCAACGSGGGGTCTGTGCACACAG	5171
Qy	5104	ccagcttgagcgaaacgacctacaccgaactgagatacctacacgctgagctatgaa	5163
Db	5170	CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGTATACCTACAGCGTGAGCATTTAGAA	5111
Qy	5164	agcgccacgcttccgaaggagaaaggcgacaggtatccggtaaagcgacgggtcgg	5223
Db	5110	AGCGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGGGTCGGA	5051
Qy	5224	acaggagagcgacaggaggagcttccagggggaaacgctggatctttatagtcctgtc	5283
Db	5050	ACAGGAGCGCACGAGGAGGCTTCAGGGGAAACGCGCTGGTATCTTTATAGTCCCTGTC	4991
Qy	5284	ggatttcgcacactcgtactcagcgtcgatttttgatgactcgtcagggggcgagc	5343
Db	4990	GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGC	4931
Qy	5344	ctatgaaaaaacgccaagcaacggcgctttttacgggttctggccttttgcggcgctttt	5403
Db	4930	CTATGAAAAAACGCCACAGCAACCGCGCTTTTACGGTTCCTGGCTTTTGGCTGGCTTTT	4871
Qy	5404	gcicacatgtcttctcgcgttatccctgattctgtgataaccgtattaccgccttt	5463
Db	4870	GCACACATGTTCTTCTCGGTTATCCCGTGATCTGTGGATAACCGTATTTACCGCGCTTT	4811
Qy	5464	gagtggactgataccgctcgcgcgacgcaacgacccgagcagcgagtcagtgagcgag	5523
Db	4810	GAGTGAGCTGATACCGCTCGCCGCGACGCCGACGACCGAGCTCAGTGAGCGAG	4751
Qy	5524	gaagcggagagcgccccaatacgaaaaacgcctctcccgcgcggttgcccgattcataa	5583
Db	4750	GAAGCGGAAGAGCGCCCAATAGCAAACGCGCTCTCCCGCGCGTGTGGCGATTTCAITAA	4691
Qy	5584	tgcagctggcacgacaggtttcccgactggaaagcgcgagtcgagcgcaacgaattaat	5643
Db	4690	TGCAGCTGGCACGACAGGTTTCCCGACTGGAAGCGGCGCAGTGAGCCGCAACCAATTAAT	4631
Qy	5644	gtgagttagctcactcattaggcaacccaggctttacactttatgcttcggggtcgtatg	5703
Db	4630	GTGAGTTAGTCTACTCATTTAGGCACCCAGGGCTTTTACACTTTATGTCTTCGGCTCGTATG	4571
Qy	5704	tgtgtggaattgtgacggataaacaatttcacagaaacacgtatgacctatgattac	5763
Db	4570	TTGTGTGGAATTGTGACGGGATAACAAATTTACACAGGAAACGCTATGACCATGATTAC	4511
Qy	5764	gcc 5766	
Db	4510	GCC 4508	

RESULT 12

US-09-194-285-7

; Sequence 7, Application US/09194285

; Patent No. 6355479

GENERAL INFORMATION:

APPLICANT: Webb, Susan R.

; APPLICANT: Winqvist, Ola

; APPLICANT: Karlsson, Lars

APPLICANT: Jackson, Michael R.

APPLICANT: Peterson, Per A.

; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems

```

; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSRI 536.1
; CURRENT APPLICATION NUMBER: US/09/194,285
; CURRENT FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4713
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-194-285-7

```

Query Match		45.5%	Score 2624.8	DB 4	Length 4713				
Best Local Similarity		99.9%	Pred. No. 0						
Matches 2626		Conservative	0	Mismatches	2				
		Indels	0	Gaps	0				
Qy	3137	cactggcgctgcttttacaa	cgctga	ctgggaaaa	ccctcg	cggttacccaacttaac	3196		
Db	2086	cactggcgctgcttttacaa	cgctga	ctgggaaaa	ccctcg	cggttacccaacttaac	2145		
Qy	3197	gccttgagcagacatccccc	ctttcg	ccagctgg	cgtaata	gagcgaagagcccg	caccgatac	3256	
Db	2146	gccttgagcagacatccccc	ctttcg	ccagctgg	cgtaata	gagcgaagagcccg	caccgatac	2205	
Qy	3257	gcccttcccaacagttgcg	agacgc	ctgtaat	gcggaat	ggcgccctgatgcgg	tattttctcc	3316	
Db	2206	gcccttcccaacagttgcg	agacgc	ctgtaat	gcggaat	ggcgccctgatgcgg	tattttctcc	2265	
Qy	3317	ttacgcatctgtcggtg	tatttaca	ccgcgcatat	ggtgc	actctcag	tacaatactgcgtctg	3376	
Db	2266	ttacgcatctgtcggtg	tatttaca	ccgcgcatat	ggtgc	actctcag	tacaatactgcgtctg	2325	
Qy	3377	atgcgcgcatagttaag	ccagccccg	acaccccg	cccaaccccg	ctgtaagcgc	cccttgacggg	3436	
Db	2326	atgcgcgcatagttaag	ccagccccg	acaccccg	cccaaccccg	ctgtaagcgc	cccttgacggg	2385	
Qy	3437	cttgtctgcctccggc	atccgctt	acagaca	agctgtg	acccgtctc	cgggagagctgcagt	3496	
Db	2386	cttgtctgcctccggc	atccgctt	acagaca	agctgtg	acccgtctc	cgggagagctgcagt	2445	
Qy	3497	gtcagaggttttcac	cgctcatc	acccgaac	cgcgagac	gaaagggcc	ctcgtgtgacgc	3556	
Db	2446	gtcagaggttttcac	cgctcatc	acccgaac	cgcgagac	gaaagggcc	ctcgtgtgacgc	2505	
Qy	3557	tattttataggttaat	gtcatat	gataata	tggtttct	ttagacgt	tcaggttggcacctttc	3616	
Db	2506	tattttataggttaat	gtcatat	gataata	tggtttct	ttagacgt	tcaggttggcacctttc	2565	
Qy	3617	ggggaaatgtcgcg	gaacccct	tattgtt	tattttc	taatacat	tccaatgtatgc	3676	
Db	2566	ggggaaatgtcgcg	gaacccct	tattgtt	tattttc	taatacat	tccaatgtatgc	2625	
Qy	3677	cgctcatgagacaata	acccctg	ataat	tgcttca	ataata	tattgaaaaagg	gaagagtaga	3736
Db	2626	cgctcatgagacaata	acccctg	ataat	tgcttca	ataata	tattgaaaaagg	gaagagtaga	2685
Qy	3737	gtattcaacattcc	ggtgc	cccttatt	cccttttt	tgccgac	tatttgcgcttccctgttt	3796	
Db	2686	gtattcaacattcc	ggtgc	cccttatt	cccttttt	tgccgac	tatttgcgcttccctgttt	2745	
Qy	3797	ttgctcaccccgaa	aacgcgt	ggtggaag	taaaag	atgctg	gaagatacag	tgttgggtgcacgag	3856
Db	2746	ttgctcaccccgaa	aacgcgt	ggtggaag	taaaag	atgctg	gaagatacag	tgttgggtgcacgag	2805
Qy	3857	tgggttacatcga	actggat	ctcaac	agcggta	agatac	cttgagag	tgttttcgccccgaag	3916
Db	2806	tgggttacatcga	actggat	ctcaac	agcggta	agatac	cttgagag	tgttttcgccccgaag	2865

Qy	3917	aacgttttccaatgatgagcaactctttaaagtcttctgctatgtggcgcgdtattatcccgta	3976
Db	2866	aacgttttccaatgatgagcaactctttaaagtcttctgctatgtggcgcgdtattatcccgta	2925
Qy	3977	ttgacgcgcggcgaagagcaactcggctgcgcgcgtatacactattctcagaatagtactggttg	4036
Db	2926	ttgacgcgcggcgaagagcaactcggctgcgcgcgtatacactattctcagaatagtactggttg	2985
Qy	4037	agtaactcacagtcacagaaaaagcattctacggatggcgtatgacagtaagagaaatacgca	4096
Db	2986	agtaactcacagtcacagaaaaagcattctacggatggcgtatgacagtaagagaaatacgca	3045
Qy	4097	gtctgcctataaccaatgagtgataacactgcgcgcgaacttactcttgacaacgatacgag	4156
Db	3046	gtctgcctataaccaatgagtgataacactgcgcgcgaacttactcttgacaacgatacgag	3105
Qy	4157	gaccgaagagctaaacgcgttttttgcacaacatagggggatcatgttaactgcgcctgatac	4216
Db	3106	gaccgaagagctaaacgcgttttttgcacaacatagggggatcatgttaactgcgcctgatac	3165
Qy	4217	gttgggaaccggagctgaatgaagccataccaaaaacgacgagcgtgacacccacgatgcctg	4276
Db	3166	gttgggaaccggagctgaatgaagccataccaaaaacgacgagcgtgacacccacgatgcctg	3225
Qy	4277	tagcaatggcaacaacgcttgcgcgaactatatacttggcgaactacttacttagcttccc	4336
Db	3226	tagcaatggcaacaacgcttgcgcgaactatatacttggcgaactacttacttagcttccc	3285
Qy	4337	ggcaacaataatagactggatggaggcggatagaagtgcaggaccactcttcgcgtcgg	4396
Db	3286	ggcaacaataatagactggatggaggcggatagaagtgcaggaccactcttcgcgtcgg	3345
Qy	4397	cccttcgcgctggctggtttatttgcgtgataaatctcggagcggctgagcgtgggtctcgcg	4456
Db	3346	cccttcgcgctggctggtttatttgcgtgataaatctcggagcggctgagcgtgggtctcgcg	3405
Qy	4457	gtatcattgcagcacctggggccagatggtaagccctccgctatcgtagttatctacacga	4516
Db	3406	gtatcattgcagcacctggggccagatggtaagccctccgctatcgtagttatctacacga	3465
Qy	4517	cggggagtcaggcaactatggtatgaacgaaaaatagacagatcgcgtgagatagggtgcctcac	4576
Db	3466	cggggagtcaggcaactatggtatgaacgaaaaatagacagatcgcgtgagatagggtgcctcac	3525
Qy	4577	tgattaagcatctggttaactcgtcacgaccaagttactcataatactttagatgatttaa	4636
Db	3526	tgattaagcatctggttaactcgtcacgaccaagttactcataatactttagatgatttaa	3585
Qy	4637	aacttcatttttaatttaaaggatctaggtagaagatcctttttgataaactcatgacca	4696
Db	3586	aacttcatttttaatttaaaggatctaggtagaagatcctttttgataaactcatgacca	3645
Qy	4697	aaatcccttaacgtgagtttgcgtccactgagcgtcagaccccgtagaaaaagatacaag	4756
Db	3646	aaatcccttaacgtgagtttgcgtccactgagcgtcagaccccgtagaaaaagatacaag	3705
Qy	4757	gatctcttgagatcctttttcttcgcgcgtaaatctgctgttcgcaacaaaaaacccac	4816
Db	3706	gatctcttgagatcctttttcttcgcgcgtaaatctgctgttcgcaacaaaaaacccac	3765
Qy	4817	cgtaccagcgggtggtttgttttgcgcggatcaagagctaccacactcttttccgaaggtaa	4876
Db	3766	cgtaccagcgggtggtttgttttgcgcggatcaagagctaccacactcttttccgaaggtaa	3825
Qy	4877	ctgggttcacagagcgcgcagataccaaaaatacttgttctcttagttagccgtagttaggcc	4936
Db	3826	ctgggttcacagagcgcgcagataccaaaaatacttgttctcttagttagccgtagttaggcc	3885
Qy	4937	accaactcaagaactctgtagcacccgcctacatacctctgcgtcctgctaatcctgttaccag	4996
Db	3886	accaactcaagaactctgtagcacccgcctacatacctctgcgtcctgctaatcctgttaccag	3945
Qy	4997	tggcgtcgtccagtggcgcgataagtcgtgtctttaccgggttggaactcaagacgatagttac	5056

```

RESULT 13
US-09-194-285-8
; Sequence 8, Application US/09194285
; Patent No. 6355479
; GENERAL INFORMATION:
; APPLICANT: Webb, Susan R.
; APPLICANT: Winqvist, Ola
; APPLICANT: Karlsson, Lars
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSRI 536.1
; CURRENT APPLICATION NUMBER: US/09/194,285
; CURRENT FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FASTSEQ for Windows Version 4.0

```


; SEQ ID NO 8
; LENGTH: 4724
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-194-285-8

Query Match 45.5%; Score 2624.8; DB 4; Length 4724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3137 cactggcgtgctttacacgctgctgactgggaaacccctggcgttaccacaaactaatc 3196
|||||
DB 2097 cactggcgtgctttacacgctgctgactgggaaacccctggcgttaccacaaactaatc 2156
OY 3197 goccctgcagcacatccccctttccgacgtggcgtaataagcgaagagcccgacccgcatc 3256
|||||
DB 2157 goccctgcagcacatccccctttccgacgtggcgtaataagcgaagagcccgacccgcatc 2216
OY 3257 goccctcccaacagctgcgacgtcgaaatggcgaaatggcgacgtgacgttatctctcc 3316
|||||
DB 2217 goccctcccaacagctgcgacgtcgaaatggcgaaatggcgacgtgacgttatctctcc 2276
OY 3317 ttaacgcatctgfcggtatttcacacccgcatatggtgcaactcagtaacatctgctctg 3376
|||||
DB 2277 ttaacgcatctgfcggtatttcacacccgcatatggtgcaactcagtaacatctgctctg 2336
OY 3377 atgcgcgatagttaagccagcccgacaccccgccacaccccgctgacgcgcccgtgacggg 3436
|||||
DB 2337 atgcgcgatagttaagccagcccgacaccccgccacaccccgctgacgcgcccgtgacggg 2396
OY 3437 ctgtgctctccggcatccgctttacagacaaagctgtgacgcgtcccgaggagctcatgt 3496
|||||
DB 2397 ctgtgctctccggcatccgctttacagacaaagctgtgacgcgtcccgaggagctcatgt 2456
OY 3497 gtcagaggtttccacgctcatccacgaaacgcgcgagacgaagggcctcgtgatacgc 3556
|||||
DB 2457 gtcagaggtttccacgctcatccacgaaacgcgcgagacgaagggcctcgtgatacgc 2516
OY 3557 tatttttataggttaatgcatataataatgggtttcttagacgtcagggtggcacttttc 3616
|||||
DB 2517 tatttttataggttaatgcatataataatgggtttcttagacgtcagggtggcacttttc 2576
OY 3617 ggggaaatgctgcgggaaacccctatttgtatttttttaataatacatcaaatatgtatc 3676
|||||
DB 2577 ggggaaatgctgcgggaaacccctatttgtatttttttaataatacatcaaatatgtatc 2636
OY 3677 cgctcatgagacaaataacccctgataaatgcttcaataatatttgaaaaggagagatga 3736
|||||
DB 2637 cgctcatgagacaaataacccctgataaatgcttcaataatatttgaaaaggagagatga 2696
OY 3737 gatttcaacatttcgctgcgcccttattcccttttttgcggcattttgccttccctgttt 3796
|||||
DB 2697 gatttcaacatttcgctgcgcccttattcccttttttgcggcattttgccttccctgttt 2756
OY 3797 ttgctcaccacgaacacgtggtgaagtaaaagatgcgtgaagatcagttgggtgacagag 3856
|||||
DB 2757 ttgctcaccacgaacacgtggtgaagtaaaagatgcgtgaagatcagttgggtgacagag 2816
OY 3857 tgggttacatcgaaactgcatccaacagcgtgaagatccttgagagttttcgccccgaag 3916
|||||
DB 2817 tgggttacatcgaaactgcatccaacagcgtgaagatccttgagagttttcgccccgaag 2876
OY 3917 aacgttttccaatgatgagcacttttaagttctctatgtggcggttattatcccgta 3976
|||||
DB 2877 aacgttttccaatgatgagcacttttaagttctctatgtggcggttattatcccgta 2936
OY 3977 ttgacgcgggcaagagaacacgtgcgcgcgatacacatattctcagaatgactgggtg 4036
|||||
DB 2937 ttgacgcgggcaagagaacacgtgcgcgcgatacacatattctcagaatgactgggtg 2996
OY 4037 agtactcaccagtcacagaaaagcatcttacggatggcgatgacagtaagagaattatgca 4096
|||||

DB 2997 agtactcaccagtcacagaaaaagcattcttaacggatggaatggaagagaaattatgca 3056
OY 4097 gtgctgccataaaccatgagtgataaacactgcggccaacttacttctgacaacgacgtggag 4156
|||||
DB 3057 gtgctgccataaaccatgagtgataaacactgcggccaacttacttctgacaacgacgtggag 3116
OY 4157 gaccgaaggagctaaaccgcttttttgacaacacatgggggagatcatgtaactcgccttgatc 4216
|||||
DB 3117 gaccgaaggagctaaaccgcttttttgacaacacatgggggagatcatgtaactcgccttgatc 3176
OY 4217 gtggaacccggagctgaatgaagccataccaaacgacgagcgtgacacacgacgtgcctg 4276
|||||
DB 3177 gtggaacccggagctgaatgaagccataccaaacgacgagcgtgacacacgacgtgcctg 3236
OY 4277 tagcaatggcaacaacgtttgcgcaactatttaactggggaactacttacttagetttccc 4336
|||||
DB 3237 tagcaatggcaacaacgtttgcgcaactatttaactggggaactacttacttagetttccc 3296
OY 4337 ggcaacaataatagactggatggaggcggaataaagtgcaggaacacacttctgcgtcgg 4396
|||||
DB 3297 ggcaacaataatagactggatggaggcggaataaagtgcaggaacacacttctgcgtcgg 3356
OY 4397 cecttccggctggtggtttattgctgataaatctggagccggtgagcgtgggtctcgcg 4456
|||||
DB 3357 cecttccggctggtggtttattgctgataaatctggagccggtgagcgtgggtctcgcg 3416
OY 4457 gtatcatgacgactggggccagatggttaagccctcccgatcgttagttatctacacga 4516
|||||
DB 3417 gtatcatgacgactggggccagatggttaagccctcccgatcgttagttatctacacga 3476
OY 4517 cgggagtcaggcaactactggtgaacgaataatagacagatcgtgagataggtgcctcac 4576
|||||
DB 3477 cgggagtcaggcaactactggtgaacgaataatagacagatcgtgagataggtgcctcac 3536
OY 4577 tgattaaagcattggttaactgctcagaccgaatttactcatatactttagattgtattaa 4636
|||||
DB 3537 tgattaaagcattggttaactgctcagaccgaatttactcatatactttagattgtattaa 3596
OY 4637 aacttcatttttaatttaaaggatctagggtgaagatcctttttgataatctcatgacca 4696
|||||
DB 3597 aacttcatttttaatttaaaggatctagggtgaagatcctttttgataatctcatgacca 3656
OY 4697 aatcccttaacgtagtgggtttcgttccactgagcgtcagaccccgtagaagaagatacgaag 4756
|||||
DB 3657 aatcccttaacgtagtgggtttcgttccactgagcgtcagaccccgtagaagaagatacgaag 3716
OY 4757 gatctcttgagatcctttttctgctgctgaatctgctgctgcaaaacaaaacacacac 4816
|||||
DB 3717 gatctcttgagatcctttttctgctgctgaatctgctgctgcaaaacaaaacacacac 3776
OY 4817 cgctaccagcgtggtttgtttgcccggatcaagagatcaacaaactcttttccgaaggttaa 4876
|||||
DB 3777 cgctaccagcgtggtttgtttgcccggatcaagagatcaacaaactcttttccgaaggttaa 3836
OY 4877 ctgggttcagcagagcgaagatacaaaaatactgctccttctagtagcgttagtgagcc 4936
|||||
DB 3837 ctgggttcagcagagcgaagatacaaaaatactgctccttctagtagcgttagtgagcc 3896
OY 4937 accacttcaagaaactctgtagcacccgctacatactcgtcctgcttaactcctgttaccag 4996
|||||
DB 3897 accacttcaagaaactctgtagcacccgctacatactcgtcctgcttaactcctgttaccag 3956
OY 4997 tggctgctgccagtggtggcgaataagtcgtgcttaccgggttggaactcaagacgatagttac 5056
|||||
DB 3957 tggctgctgccagtggtggcgaataagtcgtgcttaccgggttggaactcaagacgatagttac 4016
OY 5057 cggataaaggcagcggctcgggtgaacgggggttcgtgcacacagccacagcttgagagc 5116
|||||
DB 4017 cggataaaggcagcggctcgggtgaacgggggttcgtgcacacagccacagcttgagagc 4076
OY 5117 gaacacactacacccaagactgagatcacctacacgctgagctatgagaagaacgcacagcttc 5176
|||||
DB 4077 gaacacactacacccaagactgagatcacctacacgctgagctatgagaagaacgcacagcttc 4136

||||| 5846 gtgggaacggcgctgaatgaagaccataccaaacgacgagcgtagacaccacgatgcctg 5905
Qy 4277 tagcaatggcaacaacgctgcgcaaaactattaaactggcgaaactacttactactagcttccc 4336
Db 5906 tagcaatggcaacaacgctgcgcaaaactattaaactggcgaaactacttactactagcttccc 5965
Qy 4337 ggaacaataataagactggatgagcgagcggaataaagttagaggaaccacttctgcgctgg 4396
Db 5966 ggaacaataataagactggatgagcgagcggaataaagttagaggaaccacttctgcgctgg 6025
Qy 4397 cctctccggctggctggtttattctgataaaactggaacccggtgagcggtgggtctgcgcg 4456
Db 6026 cctctccggctggctggtttattctgataaaactggaacccggtgagcggtgggtctgcgcg 6085
Qy 4457 gtatcattgcaacactggggccagatggtgaagccctcccgatctgtagttatctacacga 4516
Db 6086 gtatcattgcaacactggggccagatggtgaagccctcccgatctgtagttatctacacga 6145
Qy 4517 cggggagtcaggcaactatggaataaagaaatagacagatcgctgagataggtgcctcac 4576
Db 6146 cggggagtcaggcaactatggaataaagaaatagacagatcgctgagataggtgcctcac 6205
Qy 4577 tgattagcattggttaactgacaccaaagtttactatatactatatacttttagattgattaa 4636
Db 6206 tgattagcattggttaactgacaccaaagtttactatatactatatacttttagattgattaa 6265
Qy 4637 aacttcatttttaattaaaggatctaggtgaagatctcttttgataaactctcatgacca 4696
Db 6266 aacttcatttttaattaaaggatctaggtgaagatctcttttgataaactctcatgacca 6325
Qy 4697 aaatcccttaacgtgagtttctgcttaactgagcgctgaagcccgtagaagaagatcaaa 4756
Db 6326 aaatcccttaacgtgagtttctgcttaactgagcgctgaagcccgtagaagaagatcaaa 6385
Qy 4757 gatctcttgagatccctttttctgctgcaatactctgctgctgcaacaaacaaacac 4816
Db 6386 gatctcttgagatccctttttctgctgcaatactctgctgctgcaacaaacaaacac 6445
Qy 4817 cgtaccagcggtggtttgttgcgagatcaagagctaccacactcttttccgaaggtaa 4876
Db 6446 cgtaccagcggtggtttgttgcgagatcaagagctaccacactcttttccgaaggtaa 6505
Qy 4877 ctggtctcagcagcagcagatcaacaaactgcttcttctagtagcgttagtaggcc 4936
Db 6506 ctggtctcagcagcagcagatcaacaaactgcttcttctagtagcgttagtaggcc 6565
Qy 4937 accactcaagaactctgtagcacgcgctacatacctcgtctgctgctgctgctgctgctg 4996
Db 6566 accactcaagaactctgtagcacgcgctacatacctcgtctgctgctgctgctgctgctg 6625
Qy 4997 tggctgctgcagtgggcgaataagtctgttcttaccggttggaactcaagacgatatgtac 5056
Db 6626 tggctgctgcagtgggcgaataagtctgttcttaccggttggaactcaagacgatatgtac 6685
Qy 5057 cggataagcgacgctggcgctgaacgggggttctgtagacacagcccgcttgagc 5116
Db 6686 cggataagcgacgctggcgctgaacgggggttctgtagacacagcccgcttgagc 6745
Qy 5117 gaacgacctacacgaactgagatcactacagcgtgagctgagctatgagaagcccgcttc 5176
Db 6746 gaacgacctacacgaactgagatcactacagcgtgagctatgagaagcccgcttc 6805
Qy 5177 ccgaaggagaaagcgcgacaggtatccggtgaagcgaggttcggaaacagagagcgca 5236
Db 6806 ccgaaggagaaagcgcgacaggtatccggtgaagcgaggttcggaaacagagagcgca 6865
Qy 5237 cgagggagctccaggggggaaacgctggtatctttagtctctgctgggttctgcgcaacc 5296
Db 6866 cgagggagctccaggggggaaacgctggtatctttagtctctgctgggttctgcgcaacc 6925
Qy 5297 tctgactgagcgtcgattttctgtagctcgctcagagggcgagcctatggaaaaacg 5356
|||||

Db 6926 tctgactgagcgtcgatgttttctgtagctcgtaagggggcgagcctatggaaaaacg 6985
Qy 5357 ccagcaacgcggcctttttacggttccctggccttttctgctggccttttctcacatgttct 5416
Db 6986 ccagcaacgcggcctttttacggttccctggccttttctgctggccttttctcacatgttct 7045
Qy 5417 tctctgcttaccctctgattctgtggataaccgctattaccgctttttagtgagctgata 5476
Db 7046 tctctgcttaccctctgattctgtggataaccgctattaccgctttttagtgagctgata 7105
Qy 5477 ccgctcgcgcagccgaacacccgagcagcagcagcagcagcagcagcagcagcagcagc 5536
Db 7106 ccgctcgcgcagccgaacacccgagcagcagcagcagcagcagcagcagcagcagcagc 7165
Qy 5537 gcccaatcagcaaacgcctctcccgcgcttggcgcgattcattaaatgcagctggcagc 5596
Db 7166 gcccaatcagcaaacgcctctcccgcgcttggcgcgattcattaaatgcagctggcagc 7225
Qy 5597 acaggtttcccgactggaaagcggcgagcagcagcagcagcagcagcagcagcagcagc 5656
Db 7226 acaggtttcccgactggaaagcggcgagcagcagcagcagcagcagcagcagcagcagc 7285
Qy 5657 ctcatagcagcccccagcgtttacactttatgcttccgctcgctatgttctgtggaattg 5716
Db 7286 ctcatagcagcccccagcgtttacactttatgcttccgctcgctatgttctgtggaattg 5745
Qy 5717 tgagcggataacaatttcacacagcaaacagcagctatgacatgattacg 5764
Db 7346 tgagcggataacaatttcacacagcaaacagcagctatgacatgattacg 7393
|||||

RESULT 15
US-08-385-335A-8
; Sequence 8, Application US/08385335A
; Patent No. 5891680
; GENERAL INFORMATION:
; APPLICANT: Lieschke, Graham J.
; APPLICANT: Mulligan, Richard C.
; TITLE OF INVENTION: Bioactive Fusion Proteins
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,335A
; FILING DATE: 08-FEB-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI95-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-385-335A-8

Query Match				45.4%; Score 2620.8; DB 2; Length 6350;			
Best Local Similarity				99.9%; Pred. No. 0;			
Matches 2633; Conservative				0; Mismatches			
				2; Indels			
				1; Gaps			
				1;			
Qy	3132	gaattcaactggcgcgtgcttttcaacagtcgtgactagggaacccctggcggttacccaact	3191				
Db	3715	GAATTCACATGGCGCTGCTTTTACAAGCTCGTGAATGGGAAACCCCTGGCGTTACCCCAACT	3774				
Qy	3192	taategccttgagcaacatcccccttccagctggcgtaataagcgaagagggccgcac	3251				
Db	3775	TAATGCCCTTGAGACATACCCCTTTGCCAGCTGGCTTAATAGGAAGAGGCCCGGCAC	3834				
Qy	3252	cgatgccttcccaacagttgcgagcgcctgaatggcggaatggcgctgatcggtatttt	3311				
Db	3835	CGATGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTGATGCGGTATT	3894				
Qy	3312	tctccttacgcacatctgcggtatttcacaccgcgcatatggtcacctctcaglacaaatc	3371				
Db	3895	TCGCCATTACGCATCTGTGGCGTATTTCACACCGCATATGGTGACATCTCAGTACAATCTG	3954				
Qy	3372	ctctgatgcgcgaatgcttaagccagccccgcgacacccgcgaacacccgcgtgaagcgccctg	3431				
Db	3955	CTCTGATCGGCATAGTTAAGCGAGCCCGACACCCGCCAACACCCGCTGAGCGGCCCTG	4014				
Qy	3432	acgggcttgctgctccccgcgcatccgcttacagaaagctgtgacogtctccgggagctg	3491				
Db	4015	ACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAAGCTGTGACCGTCTCCGGGAGCTG	4074				
Qy	3492	catgtcacagaggtttccaccgtcatccggaacgcgcga-gacgaagggccctcgta	3550				
Db	4075	CATGTGTCAGAGGTTTTCACCGTCAATCCGGAACCGCGGATGACGAAGGGCCCTCGTA	4134				
Qy	3551	taogccatttttatagggttaatgcatgataataatggtttctcttagacgtcaggtggca	3610				
Db	4135	TACGCCATTATTTATAGTTAATGTCTATATAATAATGTTTCTTAGACGTGAGTGGCA	4194				
Qy	3611	cttttcggggaatgctgcgggaacccctatttgtttatttttcttaataacattcaata	3670				
Db	4195	CTTTTCGGGGAATGTGCGCGGAACCCCTATTGTTTATTTTCTTAATACATTTCAAAATA	4254				
Qy	3671	tgatccgcgtcatagacaataaccctgataaatgcttcaataataattgaaaaggaaga	3730				
Db	4255	TGTATCCGCTCATGACACAATAACCTGATTAATGCTTCAATAATATTGAAAAGGAAGA	4314				
Qy	3731	gtatgagtatcaaacatttcogtgcgccttattcccttttttggcggaattttgccttc	3790				
Db	4315	GTATGAGTATTCAACATTTCCGCTGTCGCCCTTATTCCTCTTTTGGCGCATTTTGCCCTTC	4374				
Qy	3791	ctggttttgcctccacagaacgcgtgtaagtaaaagatgctgaagatcaggttgggtg	3850				
Db	4375	CTGTTTTTGTCAACCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG	4434				
Qy	3851	cacgagtggtttacatcgaactggatctcaacagcggtgaagatcccttgagagttttcgcc	3910				
Db	4435	CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTGAAGATCTTGAGAGTTTCGCC	4494				
Qy	3911	cgaagaaacggttttccaaatgatagcaccttttaaagttctgctatgtgcggtattat	3970				
Db	4495	CGAAGAAGGTTTTTCAATGATGAGCACATTTAAAGTTCTGCTATGTGCGCGGTATTAT	4554				
Qy	3971	ccggtattgacgggggaagcaactcgtgcgcgatacaactattctcagaatgact	4030				
Db	4555	CCCGTATTGACGGCGGGAAGAGCAACTCGGTTCGGCCGCATACACTATTCTCAGAATGACT	4614				
Qy	4031	tggttgactactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaat	4090				
Db	4615	TGTTTCAGTACTCACCCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA	4674				
Qy	4091	tatgcagtgctgccataacacatagtgataaacactgcggccaaacttactcttgacaacga	4150				
Db	4675	TATGCAGTGTGCTCATTAACCATGAGTGATAACACTCGCGGCCAACTTACTTCTGACAACGA	4734				
Qy	4151	tcggaggaccgaagagctcaacgcgttttttgcacaaacatgggggcatgtaactcgcc	4210				

```
Db 5815 AGCGCAGGAGGAGTCCAGGGGAAACGCCCTGGTATCTTTATAGTCCTGCTCGGGTTTC 5874
Qy 5291 gccacctgtacttgagcgtcgatcttttctgctcagggggcgagccctatgga 5350
Db 5875 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA 5934
Qy 5351 aaaacgcagcaacgcgcctttttacggtttccctggtcttttctgctggtcttttgcctca 5410
Db 5935 AAAAGCGCAGCAACGCGGCCCTTTTACGGTTCCCTGGCCCTTTTGTGGCCCTTTTGTCTACA 5994
Qy 5411 tgttcttctcgtcgttaccctctgattctgtggataaccgtattacccgcttttgcgtgag 5470
Db 5995 TGTTCCTTTCCCTGCGTTATCCCTGATTCCTGTGGATAACCGTATTACCGCCCTTTGAGTGAG 6054
Qy 5471 ctgataccgctcgcgcgagccgagccgagcagcagcagcagcagcagcagcagcagcagcag 5530
Db 6055 CTGATACCGCTTCGCGCGCAGCGAGCCGAGCCGAGCGAGCGAGTCAGTGAGCGAGGAAGCGG 6114
Qy 5531 aagagcgcacaaacgcgcctctcccgcgcttggcgattcattcaatgcagct 5590
Db 6115 AAGAGCGCCCAATACGCANACCGCTCTCCCGCGCTTGGCCGATTCAATATGCAGCT 6174
Qy 5591 ggcacgacaggtttcccgactggaaagcggcagtgagcgcaacgcaattaatgtgagtt 5650
Db 6175 GGCAGCAGAGTTCCTCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTT 6234
Qy 5651 agctcaactcattaggaccccgaggtttacactttatgcttcgggctcgatgtgtgtg 5710
Db 6235 AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 6294
Qy 5711 gaattgtgagcgataaacaatttcacaggaacagcctatgaccatgattacgcc 5766
Db 6295 GAATTGTGAGCGGATACAAATTTACACAGGAAACAGCTATGACCATGATTACGCC 6350
```

Search completed: August 31, 2002, 15:34:03
Job time: 6805 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:10:48 ; Search time 9324.87 Seconds
(without alignments)
3871.183 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgagggcccccagtgctct.....aattgctcagccacctga 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1725	100.0	1725	6	AX275256
2	1722.4	99.8	3096	6	AR070206
3	1722	99.8	2218	9	HUMACHE
4	1722	99.8	2256	6	AR070205
5	1722	99.8	3016	6	AR070207
6	1718.6	99.6	5767	6	AX275254
7	1718.6	99.6	14446	6	AX275255
8	1364.8	79.1	1947	4	OCU05036
9	1282	74.3	2089	6	AX306139
10	1282	74.3	2089	10	MMACHE
11	1276.2	74.0	2066	10	S50879
12	1198	69.4	4185	9	HUMACHEB
13	1198	69.4	172358	9	AC011895
14	1198	69.4	194681	2	AC084057
15	1126	65.3	34921	9	HSFA002993
16	1070	62.0	1355	9	HUMACHE01
17	890.8	51.6	3958	4	AF053485
18	803.6	46.6	984	4	BTACHE1
19	775	44.9	296820	10	AF312033
20	741.8	43.0	1698	6	AX019953
21	741.8	43.0	2025	5	BFU54591
22	741.8	43.0	2478	6	AX019954
23	566.2	32.8	2490	5	TCACER
24	553.4	32.1	2357	5	TWACHE
25	496	28.8	4472	5	AF030422
26	490.4	28.4	779	9	HUMACHE02
27	482	27.9	13481	5	DRE251640
28	431	25.0	503	4	BTACHE2
29	404	23.4	1922	4	OCBCHX1
30	394	22.8	1809	4	AF178685
31	390.8	22.7	1932	4	AF053483
32	389.2	22.6	1931	4	AF053484
33	387.8	22.5	1971	5	GGA306928
34	380	22.0	2173	9	BC018141
35	380	22.0	2381	6	AR070209
36	380	22.0	2381	9	HUMCHEF
37	380	22.0	2416	6	AR070208
38	380	22.0	2416	6	AR169579
39	380	22.0	2416	6	AR169580
40	380	22.0	2416	9	HUMCHEB
41	378.6	21.9	3600	3	BFU74380
42	371.4	21.5	2682	9	HUMCHEBG2
43	371.4	21.5	171083	2	AC009811
44	371.4	21.5	186526	2	AC048332
45	370	21.4	1878	10	AF244349

ALIGNMENTS

RESULT	1
AX275256	
LOCUS	AX275256 1725 bp DNA linear PAT 29-OCT-2001
DEFINITION	Sequence 5 from Patent WO0171014.
ACCESSION	AX275256
VERSION	AX275256.1 GI:16547676
KEYWORDS	synthetic construct.
SOURCE	synthetic construct
ORGANISM	artificial sequence.
REFERENCE	1 (sites)
AUTHORS	Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE	Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL	Patent: WO 0171014-A 5 27-SEP-2001; BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor, Tsairir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)
FEATURES	Location/Qualifiers
source	1..1725
	/organism="synthetic construct"
	/db_xref="taxon:32630"
	/note="synthetic human acetylcholinesterase gene optimized"

BASE COUNT	273 a	553 c	571 g	328 t
ORIGIN	for expression in plants"			

Query Match 100.0%; Score 1725; DB 6; Length 1725;
Best Local Similarity 100.0%; Pred. No. 1e-294;
Matches 1725: Conservative 0; Mismatch 0; Indels 0;

Qy	1	atgagggccccgagtgctgctgacacgaccttccctggcttccccactcttctctc	60
Db	1	ATGAGGGCCCCGAGTGTCTGTCTGTACACAGCCCTCCCTGGCTTCCCACTTCTCTCTCT	60
Qy	61	ctctcttggtctctgggttgaggagtggggcttgaggccgagagatgcagagctgctg	120
Db	61	CTCTCTTGGCTTCCCTGGGTGAGGAGTGGGGGTGAGGGCGGAGGATGCAGAGCTGCTG	120
Qy	121	gtgacgggtgctggggggccggctgctgggggctatcgccctgaagaccctccggggccctgtc	180
Db	121	GTGACGGTGGCTGGGGGGCGGCTGCGGGGGCTATTCGCTTGAGACCCCGGGGGCCCTGTCT	180
Qy	181	tctgctttctctgggcatccctcttgggagccacccatgggacccctgcgtttcttgcca	240
Db	181	TCTGCTTTCCTGGGCATCCCTTTTCGAGGACCACCCATGGACCCCGTCTCTTCTTGCCA	240
Qy	241	ccgagagcccaagcaagcttggtcagggttgtagagcgtacaaccttccaagatgtctgc	300
Db	241	CCGAGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGCGCTACAACCTTCCAAGATGTCTGC	300
Qy	301	taccaatatgtggacacccctatacccaagttttaggggaccagagatgtggaaccccaac	360
Db	301	TACCAATAATGTGGACACCCCTATACCCAGGTTTTTAGGGGCACCCAGAGATGTGGAACCCCAAC	360
Qy	361	cgtgagctgagcaggagactgcctgtacctcaacgctgtggacacctacccccggcctaca	420
Db	361	CGTGAGCTGAGCAGGAGACTGCCTGTACTCAACGTGTGGACACCATACCCCGGCTTACA	420
Qy	421	tcccccaacctgtctctgtctggatctatgggggttgcttctacagtgagggcctctcc	480
Db	421	TCCCCCAACCCCTGCTCTGTCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC	480
Qy	481	tggacgtgtacgattggccgctctcttggtacagcgagagagactgtgctgggtggtccatg	540
Db	481	TTGGACCTGTACGATGGCCGCTTCTTGGTAGACGCCGAGAGGACTGTGCTGGTGTCCATG	540
Qy	541	aactaccgggttgggagcctttggcttctctggcctgcgggggagccgagagccccgggc	600
Db	541	AACTACCGGGTGGAGGCTTTGGCTTCTGCGCCTGTCGCGGGAGCCGAGAGGCCCGCGGC	600
Qy	601	aatgtgggtctcttgatcagagagcttgccctcagtggttgaggagaaacttggcagcc	660
Db	601	AATGTGGGTCTCTTGGATCAGAGCTTGCCCTCTACGTGGGTGAGAGAACGTGGCAGCC	660
Qy	661	ttcgggggtgacccgacatcagtgacgctgtttgggagagccgagagccgctcggtg	720
Db	661	TTCCGGGGTGACCCGACATCAGTGCACGCTGTTTGGGAGAGCCCGGGAGCCGCGCTCGGTG	720
Qy	721	ggcatgcacctgtctcccccgcagcgggggctgttccacgggcccgtgctgcagagc	780
Db	721	GGCATGCACCTGTCTCCCGCCAGCGGGGCTGTGTTCCACAGGSCCTGTCTGCAGAGC	780
Qy	781	ggtgcccccaatggacacctgggccacggtgggcatgggagagcccgctcgaggggccag	840
Db	781	GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGATGGGAGAGGCCCGTCTGCAGGGCCAG	840
Qy	841	cagctggcccaaccttgggctgttcctccagcgagcactggtgggaatcacacagagctg	900
Db	841	CAGCTGGCCCACTTGTGGGCTGTCTCTCCAGGGGCACTGCTGGGGAATGACACAGAGCTG	900
Qy	901	gtagcctgccttcgggacacgaccagcgaggtcctgggtgaacccacgaattggcactgtctg	960
Db	901	GTAGCCTGCCTTCGGACACGACGACGCGAGGTCTGGTGAACCCACGAATGGCACGTGCTG	960

Qy	961	cctcaagaagacgctcttcgggttctccttgctgcctgtggttagatgaggagacttoctcagt	1020
Db	961	CCTCAAGAAACGCTTCTCGGGTTCTCCTTTCGTGCCTGTGGTAGATGGAGACTTCTCTCACT	1020
Qy	1021	gaacccacagagccctcatcaacgcgagagactttccaagcgctcaagtgctggtgggt	1080
Db	1021	GACACCACGAGGCCCTCATCAACGGCGGAGACTTCCACGGCCTCGAGGTGCTGGTGGGT	1080
Qy	1081	gtggtgaagtagtagggctcgtaatttctcgttgttacggggcccceagggcttcagcaaaagac	1140
Db	1081	GTGCTGAAGGATCAGGGCTCGTATTTCCTTGCTTTCGGTTTTCGGGGCCCCAGGCTTCAGCAAAAGAC	1140
Qy	1141	aacgagtctctcatcacgcgccccgaggttctctggccggggtgcggggttcggggttccccag	1200
Db	1141	AACGAGTCTCTCATCAGCCGGGCCGAGTTCCTGGCCGGGGTGGCGGTCTCGGGTTCGCCAG	1200
Qy	1201	gtaagtgaacctgcagccgagagctgtggtctctgcattaacacagaactggtcgtcatccagag	1260
Db	1201	GTAAGTAGACTTGGCAGCCGAGGCTGTGGTCTCTGCATTTACACAGACTGGCTGCTCATCCCAG	1260
Qy	1261	gaccggcagcctgagggaggcccttgagcgtgtggtggcgaccacaatgtcgtgtgc	1320
Db	1261	GACCCGGCACCCCTGAGGGAGGCCCTGAGCGATGTGTGGCGGCACACAAATGTCGTGTGC	1320
Qy	1321	cccggtgccagctggtcgtgggcagctggctgccagggtgcccggggtctacgcctacgtc	1380
Db	1321	CCCGTGGCCACAGTGGCTGGCGGAC'TGGCTGCCAGGGTGC'CCGGGTCTACGCGCTACGTC	1380
Qy	1381	ttagaacacocgtctctccacgctctcctctggcccccctgtgga'tgggggtgccccacggctac	1440
Db	1381	TTTGAACACCGCTCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGGTGCCCCACGGGTAC	1440
Qy	1441	gaqatcgagttcatcttfggatcccccttgaccctctcgaatactacacggcagagagag	1500
Db	1441	GAGATCAGAGTTCACTTTTGGATCCCCCTTGGACCCCTCTTCGAAAATACACGGCAGAGAG	1500
Qy	1501	aaaatcttgcacagcagctgatgcgatactgtggccaaactttgcccgcacaggggatcccc	1560
Db	1501	AAATCTTTCGCCAGGACTGATGCGATACTGGGCCAATTTTGGCCGACACAGGGGATCCC	1560
Qy	1561	aatgagccccagacccccaaaggccccacaatatggccccccgtacacggcgggggtcctcagcag	1620
Db	1561	AATGAGCCCCGAGACCCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGGCTCAGCAG	1620
Qy	1621	tacgttagctctggaacctgcgcgcctgtaggtgtcgcgcggggggtcgcgcgccagagcctgc	1680
Db	1621	TACGTTAGTCTGCACTTGGCGGCGCTGGAGGTGGCGGGGGGCTCGCGGCCCCACAGCCCTGC	1680
Qy	1681	gcctcttggaacgccttctccccaaattgctcagcgccacctga	1725
Db	1681	GCCTTCTGGAACCGCTTCTTCCCAAATTTGCTCAGCGGCCACCTGA	1725

RESULT 2

AR070206	AR070206	3096 bp	DNA	linear	PAT 18-FEB-2000
LOCUS					
DEFINITION	Sequence 6 from patent US 5891725,				
ACCESSION	AR070206				
VERSION	AR070206.1	GI:7221094			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 3096)				
TITLE	Soreg, H., Zakut, H. and Eckstein, F.				
JOURNAL	Synthetic antisense oligodeoxynucleotides and pharmaceutical				
FEATURES	compositions containing them				
source	Patent: US 5891725-A 6 06-APR-1999;				
	Location/Qualifiers				
	1. .3096				
	/organism="unknown"				
BASE COUNT	509 a 1089 c 872 g	626 t			
ORIGIN					

Query Match 99.88; Score 1722.4; DB 6; Length 3096;
Best Local Similarity 99.9%; Pred. No. 2.7e-294;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	atgagggcccccagtgctactacacacgcttccctggcttccceactcttctc	60
Db	160	ATGAGGCCCCCGCAGTGTCTGTCACACGCTTCCCTGGCTTCCCACTCCTTCTC	219
Qy	61	ctctctgtctccctgggtggagagtgaggctgagggccggagatgcagagctgtg	120
Db	220	CTCCTCTGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGAGATGCAGAGCTGTG	279
Qy	121	gtacaggtgtcgtggggcgggctcgtgggcatctgcctgaagacccccggggccctgtc	180
Db	280	GTGACGGTGTGGGGGGCGGCTGCGGGGCATTTGCGCTTGAAGACCCCGGGGCGCTGT	339
Qy	181	tctgttctctgggcatcccttgcgagaccacccatggagaccctgcgtttctgcga	240
Db	340	TCTGCTTCTTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCGTCTGTGCCA	399
Qy	241	ccgagagcccaagcagccttgggtcaggggtgtgagcgtctaacacttccagagtgtgc	300
Db	400	CCGAGGCCCAAGCAGCCTTGTGAGGGTGTGAGACGCTACAACCTTCCAGAGTGTCTGC	459
Qy	301	taccaatgtgagacacccctataccaggttttgagggcaccagagatgtgaaccccaac	360
Db	460	TACCAATATGTGGACACCCCTATACCAGGTTTTCAGGGCACCGAGATGTGGAACCCCAAC	519
Qy	361	cgtgagctgagcagagactgctgtacctaactgaactgtgagaccatacccccgcctaca	420
Db	520	CGTGACTGACGAGAGTGTGCTGTACCTCAACGTGTGGACACCATACCCCGGGCTACA	579
Qy	421	tccccacccctgtctcgtgtgatctatgggggtggtcttctacagtggggccctctcc	480
Db	580	TCCCCACCCCTGCTCTGTGATCATGGGGTGTCTTACAGTGGGGCTCTCTCC	639
Qy	481	ttgagctgtacagatggcgtcttctgtacagccgagagagactgtctggttccatg	540
Db	640	TTGGAAGTGTACGATGGCGGCTTCTGTGTACAGGCGGAGAGAGTGTCTGTGTCCATG	699
Qy	541	aactaccgggtggagaccttggcttctgctccctgcggggagccagagagcccgagc	600
Db	700	AACATACCGGGTGGAGACCTTTGGCTTCTGCTTCCCGGGAGCCGAGAGCCCGGGC	759
Qy	601	aatgtgggtctcctggatcagaggtggtgcccctgagtggtgtgaggaacgtggagacc	660
Db	760	AATGTGGGTCTCCTGGATCAGAGGTGSCCTGCAGTGGGTGCAGGAGACGTGGCAGCC	819
Qy	661	tctgggggtgaccgacatacagtgacgtgttgggggagcgcgggagccctcggtg	720
Db	820	TTCCGGGGGTGACCCACATCAGTACGCTGTTTGGGGAGAGCGGGAGCCCTCGGTG	879
Qy	721	ggcatgacactgtctcccccagcgggctgttcccacagggccgtactgcagagc	780
Db	880	GGCATGACACTGTCTCCCGCCAGCCGGGGCTGTTCACAGGGCCGTCTGCAGAGC	939
Qy	781	gggtgcccccaatggaacctgggacagtggtggcatgggagagggccgtcgcagggccacg	840
Db	940	GGTGGCCCAATGGACCTTGGGCCACGCTGGCATGGCATGGGAGAGGCCCTCGCAGGGCCACG	999
Qy	841	cagctgggcccacctgtgggtgtctctccagcgggcaactggtgggaatgacacagagctg	900
Db	1000	CAGCTGGCCCCACCTTGTGGGTGTCTCCAGCGGCACTGTGTGGGAATTCACACAGAGCTG	1059
Qy	901	gtagcctgcttcggacacgaccagcagctcctggtggaaccagaaatgcacgtgtg	960
Db	1060	GTAGCCTGCTTCGGACACAGACCGCAGGTCTCTGTGTGAACCAAGAAATGGACGCTGTG	1119
Qy	961	cctcaagaaagcgtcttcctggtctctctcgtgctgtggttagatggagacttctcag	1020
Db	1120	CCTCAAGAAAGCGCTTCCCGGTCTCTCTCTGCTGTGTGTAGATGGAGACTTCTCTCAGT	1179

Qy	1021	gacacccccagaggccctcatcaacgctgggagacttccacggcctgcaggtgtgtggt	1080
Db	1180	GACACCCAGAGGCCCTCATCAACGCGGAGACTTCCACGGCCTGCAGGTGTGTGGGT	1239
Qy	1081	gtggtgaagagatgaggtcgtattttctgtttacggggggcccccaggttcaagaagac	1140
Db	1240	GTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGGGCCAGGCTTCAGCAAGAC	1299
Qy	1141	aacgagttctctcatcagcggggccaggttccctggcgggggtgcgggttcgggttccccaag	1200
Db	1300	AACGAGTCTCTCATCAGCGGGCCGAGTTCCTGGCGGGGGTGCGGGTTCGGGTTCCCCAG	1359
Qy	1201	gtaagtgaacctggcagcggaggtgtgtcctgcattacacagactggtgcattccccag	1260
Db	1360	GTAAGTACCTTGGCAGCGGAGGCTGTGCTCTGTCATTACACAGACTGGTGTGATCCCGAG	1419
Qy	1261	gaccggcacacccctgagggagccctgagcagatgtggtggggagaccacaatctcgtgtgc	1320
Db	1420	GACCCGGCAGCGCTGAGGGAGGCCCTGAGCGATGTGGTGGGGGACCAACAATGTCTGTGTG	1479
Qy	1321	cccgtggccccagctggtggggcagctggtgcccaggggtgcccgggttctacgctacgtc	1380
Db	1480	CCCGTGGCCCCAGCTGGCTGGGGACTGCTGCTGCCAGGGTGCCTGGGTCTACGCTACGTC	1539
Qy	1381	tttgaaacacctgtctccacgctctcctggccccctgtgagtggtgggtgccccacaggtac	1440
Db	1540	TTTGAACACCGTGTCTCCACGCTCTCTGGCCCTGTGGATGGGGGTGCCCCACGCTAC	1599
Qy	1441	gagactgagttcatcttgggagccccctggacccctcgaacactacacggcagaggag	1500
Db	1600	GAGATCGAGTTTCATCTTTGGGATCCCTTGGACCCCTCTCGAAACTACACGGCAGAGGAG	1659
Qy	1501	aaaacttctgcccagcagctgactgactgactggtggcgaactttggccgcacagggatccc	1560
Db	1660	AAATCTTTCGCCACGACTGATGGATACTTGGGCCAATTTGCCCGGCACAGAGGGATCCC	1719
Qy	1561	aatgagccccagagaccccccaagcccccaatggcccccgctacacgctgggggtcagcag	1620
Db	1720	AATGAGCCCCGAGAGACCCCAAGGCCCAATGGCCCGGTACACGGCGGGGCTCAGCAG	1779
Qy	1621	tacgttagcttggaacctgcccgcctgaggtgaggtgcccgggggctgcggccccagcctgc	1680
Db	1780	TACGTGTGTGTGAGCTGCGGCCCTGTGAGGTGTGGGGGGGCTGCGGGCCAGGCGCTGC	1839
Qy	1681	gocctctggaacgcttctctcccaaatgtcctcagcgccacactg	1724
Db	1840	GCTTCTGGACCGCTTCTTCCCCAAATGCTCAGCGGCCACCGG	1883

RESULT	3
HUMACHE	
LOCUS	2218 bp mRNA linear PRI 09-JUN-1995
DEFINITION	Human acetylcholinesterase (ACHE) mRNA, complete cds.
ACCESSION	M5040
VERSION	M5040.1 GI:177974
KEYWORDS	acetylcholinesterase.
SOURCE	Human 21-week old fetus DNA, and cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1. .2218
TITLE	Location/Qualifiers
JOURNAL	1. .2218
MEDLINE	/organism="Homo sapiens"
FEATURES	/db_xref="taxon:9606"
SOURCE	

```
/map="7q22"
/dev_stage="21-week-old fetus"
/tissue_lib="lambda-gt10"
1. .2218
/gene="ACHE"
/note="G00-118-746"
/product="acetylcholinesterase"
1. .2218
/gene="ACHE"
157. .2001
/gene="ACHE"
/EC_number="3.1.1.7"
/codon_start=1
/product="acetylcholinesterase"
/protein_id="AAA68151.1"
/db_xref="GI:177975"
/db_xref="GDB:G00-118-746"
/translation="MRPQCLHTPSLASPLLLLLLLGGVGAGREDAELLYTVR
GRLRGRLKTPGGVSALFGLFPAEPMPRRFLPPPKQPSGVYVDATTFQSYCYQ
YDTLYPFGTEGEMNPNRESEDCLYLNWTPYPRPTPLVIMVIGGGYSYGS
LDVYDRELVQAERTLVSMNYRVGAFGLFALPGREAPGNVGLLDORLALQVQENY
AAFQGDPTSVTLFEGSSAGAAVSVMHLLSPPSRGLFRAVLQSGAPNGPWATVGMGEAR
RRATLAHLVGCPCPGTGGNDTELVACLRTPAQVLVNHWHVLPQESVFSPFVV
DGFDFLSDPEALINAGDEHGLQVLGVYKDEGSYELVYGAPGFSKDNESLISRAELA
GVRVGPQVSDUAEAAYLVHYDNLHPEDPARREALSDVYGDHNVCPVQAQAGELA
AGARVYAYVFHRASTUSWPLMGVPVPHGYEIEFIFGIPLDPSRNTAEKIFPAQRLLM
RYWANFARTGDPNEPRDPKAPQMPVPTAGAQOYVSLDLRLPLEVRRLRAQACAFWNR
LPKLLSADTDTLDEAERQWKAEPHRWSSVMVHMKQFDHYSKQDRCSDL"
229. .291
sig_peptide
/gene="ACHE"
/note="G00-118-746; does not fit consensus"
310. .1998
/mat_peptide
/gene="ACHE"
/EC_number="3.1.1.7"
/note="G00-118-746"
/product="acetylcholinesterase"
BASE COUNT 355 a 755 c 680 g 428 t
ORIGIN

Query Match 99.8%; Score 1722; DB 9; Length 2218;
Best Local Similarity 100.0%; Pred. No. 3.4e-294;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagcccccagtgctgtgcacagccttcctggctcccccacatctcttcctc 60
Db 157 ATGAGCCCCCAGTGTCTGTGCACAGCCCTTCCTGGCTTCCCGACTCTCTCTC 216

QY 61 ctccctggctcctggtagagagtgagggtgagggcgagggcgagagatgcagagctg 120
Db 217 CTCCTCTGGCTCTCTGGTGGAGAGTGGGGGCTGAGGGCGGGAGATGCAGAGTGC 276

QY 121 gtaacggtgctggggcggtgctggggcattcgctgaagacccccggggcgctg 180
Db 277 GTGACGGTGGTGGGGCGGCTGCGGGGCATTCGGCTGAAGACCCCCGGGGCCCTG 336

QY 181 tctgcttctgggacatcccccttgcggagcaccatgggaccccgcttcttgcca 240
Db 337 TCTGCTTCTCTGGGCATCCCTTTGGGAGCCACCCCATGGACCCCGCTGCTTCTGCCA 396

QY 241 ccgagacccaagacgcttgtaggggtgtagagcgtacacattccagagtgctgc 300
Db 397 CCGAGCCCAAGCAGCGCTTGGTCAGGGGTGATAGCGCTACAACTTCCAGAGTGTGC 456

QY 301 taccatagtgacaccctaaccaggttttgagggcacccagagatgtggaaccccaac 360
Db 457 TACCATATGTGACACCCCTATACCCAGGTTTGGAGGCCACCGAGATGTGAACCCCAAC 516

QY 361 cgtgagctgagcaggaactcgtctaccccaacgctgtgacaccatacccccggtcaca 420
Db 517 CGTGAAGTGAAGGAGTCCCTGTACCTCAAGCTGTGGACACCATACCCCGGCTACA 576

QY 421 tcccccaacccctgctcgtatgtagctatgagggtgggttctacagtgaggcctctcc 480
```

```
Db 577 TCCCCCACCCTCTCTCTGCTCGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCTCC 636
QY 481 ttgagcgtgtacgatggcgcttcttggtacagggccgagaggaactgtgctgggtgctcatg 540
Db 637 TTGGAGCTGTACGATGCCGCTTCTTGGTACAGCCGAGAGGACTGTGCTGGTGTCCATG 696
QY 541 aactaccgggtggagcccttggcttccctggccctgcccggggagcagagagccccgggc 600
Db 697 AACTACGGGTGGAGCCTTTGGCTTCTTGGCCCTCGGGGAGCCGAGAGCCCGGGC 756
QY 601 aatgtgggtctcctggatcagagagctggccctcagtgagggtcaggagaaactgagcagcc 660
Db 757 AATGTGGGTCTCTGTGATCAGAGGCTGGCCCTCAGTGGGTGGTGGAGAACGTTGGCAGCC 816
QY 661 ttcgggggtgacccgcacatcagtgacgctgtttggggagagcgggagcgccctcggt 720
Db 817 TTCGGGGGTGACCCGACATCAGTGCCTGTTGGGAGAGCGCGGGAGCCGCTCGGTC 876
QY 721 ggcattgacatgctgtcccccgcacagccgggctgtccacagggccgtgctgacagagc 780
Db 877 GGCATGCACTGTCTGCCCGCCAGCCGGGGCTGTCTTCCACAGGGCCGCTGCTGCAGAGC 936
QY 781 ggtgcccccaatgagccctgggcccacaggtggcatgggagagggccgtgcaggggccaag 840
Db 937 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGCATGGGAGAGCCCGCTCGAGGGCCACG 996
QY 841 cagctggcccaaccttgggtgtctccagcgcgctggtgggaatgacacagagctg 900
Db 997 CAGCTGGCCCACTTGTGGGCTGTCTCCAGGGCGGCACTGTGGGGAATGACACAGAGCTG 1056
QY 901 gtacgtcgtctcgagacagcagcgaggtcctggtgaacacacagtgagcagtgctg 960
Db 1057 GTAGCCTGCTTCGGACAGACACAGCGCAGTCTGTGTGAACACCAATGGCACGCTGCTG 1116
QY 961 cctcaagaaagcgtctccgggttctcctctgctgctgtagatgagagacttccctcagt 1020
Db 1117 CCTCAAGAAAGCGCTTCCGGTTCCTCTGCTGCTGTGGTGTAGATGGAGACTTCTCTCAGT 1176
QY 1021 gacacccagagggcctcatcaacgcgagagcttccagcgctgcaggtgctgggtgggt 1080
Db 1177 GACACCCAGAGGCCCTCATCAACGCGGGAGACTTCCAGGCCCTGCAGGTGCTGGTGGT 1236
QY 1081 ggtgtaagagtaggggctctgtatttctggtttacggggcccccaggtctcagcaagac 1140
Db 1237 GTGGTGAAGGATGAGGCTCGTATTTCTGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1296
QY 1141 aacgagctctcatcagccggcgaggttcttgccggaggtgcggtcggggttccccag 1200
Db 1297 AACGAGTCTCTATACGCCGGCGGAGTTCTTGCCCGGGGTGCGGGTTCGCCAG 1356
QY 1201 gtaagtgaacctggcagcgaggtggtggtctgctgattacacagactggctgcatccccag 1260
Db 1357 GTAAGTGACTGGCAGCGCGAGGCTGTGGTCTTCATTACACAGACTGGCTGCATCCCGAG 1416
QY 1261 gacccggacgctgaggagggccctgagcgaatggtgggggacacacaaatgctgctg 1320
Db 1417 GACCCGGCAGCGCTGAGGAGGGCCCTGAGCGATGTGTGGGGCCACCAATGTCTGTGTC 1476
QY 1321 cccgtgagccagctgctgggagactggtgccaggtgcccgggtctctacgctcagtc 1380
Db 1477 CCGTGGCCAGCTGGCTGGCGGACGTGGCTGCCCGGGTCCCGGGTCTACGGCTACGTC 1536
QY 1381 ttgaaacacgctgcttccacgctcctcctggcccttggatgggggtgccccacgctac 1440
Db 1537 TTTGAACACCGCTTCTCCAGCTCTCTTGSCCCCTGTGGATGGGGTGCCTCCAGGCTAC 1596
QY 1441 gagatcagtgatcttctgggatccccctggacccctctctgaacactacacgagggag 1500
Db 1597 GAGATCAGTTCATCTTGGGATCCCGCTGGACCCCTCTCGAAACTACACGGGAGAGGAG 1656
QY 1501 aaaaattcggccagcagctgtagctactgagcactgagcacttgcgcgacaggggagtc 1560
```

Db 1657 AAAATCTTGGCCAGCGACTGATCGATACCTGGCCAACTTTTGCCCGCACAGGGGATCCC 1716
Qy 1561 aatagagcccgagagcccaagagcccaacaaatgcccccgatcacagcgggggctcagcag 1620
Db 1717 AATGAGCCCGAGAGACCCCAAGGCCCAACAAATGGCCCGCTACAGCGGGGGCTCAGCAG 1776
Qy 1621 tacgttagctgagacctgagccgctgagagtgcggggggctgcgcgcccagggcctgc 1680
Db 1777 TACGTTAGCTGAGACCTGCGGCCCTGGAGGTGGGGGGGCTGGCGGCCCCAGGCCTGC 1836
Qy 1681 gcctctggaaccgcttctctctcccaaaatgtctcagcgccacc 1722
Db 1837 GCCTTCTGGAACCGCTTCTCTCCCAAAATGCTCAGCGGCACC 1878

RESULT 4
AR070205 AR070205 2256 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 5 from patent US 5891725.
DEFINITION AR070205
ACCESSION AR070205
VERSION AR070205.1 GI:7221093
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2256)
AUTHORS Soreg,H., Zakut,H. and Eckstein,F.
TITLE Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them
JOURNAL Patent: US 5891725-A 5 06-Apr-1999;
FEATURES Location/Qualifiers
source 1..2256
BASE COUNT 390 a 757 c 680 g 429 t

Query Match 99.88; Score 1722; DB 6; Length 2256;
Best Local Similarity 100.0%; Pred. No. 3.4e-294;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagggcccgagtgctgtgcacagccttccctggcttcccactctctctc 60
Db 160 ATGAGGCCCCGCGAGTGCTCTGCACAGCGCTTCCCTGGCTTCCCACTCTCTCTC 219
Qy 61 ctctcttgctctcgggtgagggagtggggctgagggcgaggatgcagagctgctg 120
Db 220 CTCTCTGGCTTCTGGGTGGAGAGTGAGGGGCTGAGGGCGGGAGATGCAGAGCTGCTG 279
Qy 121 gtgacgtgctggggcggtcggtggggcattcgctgaagacccccgggggcccctgtc 180
Db 280 GTGACGGTGCCTGGGGCGGGCTCGGGGGCATTCGCTGGAAGACCCCGGGGGCCCTGTG 339
Qy 181 tctgctttctgggcatccctttgagagccaccatgggagcccgctcgctttctgcga 240
Db 340 TCTGCTTCTTGGGCATCCCTTTGGGAGCCACCCATGGACCCCGCTGCTTCTGCGCA 399
Qy 241 ccgagagcccaagcagccttgctcaggggtgtagacgtacaaaccttcagagtgctgc 300
Db 400 CCGGAGCCCAAGCAGCCTTGGTCAGAGGGGTGTAGACGCTTACAACCTTCCAGAGTGTCTGC 459
Qy 301 taccaatatgggacacctataccaggtttttagggcaccagagatgtgaaacccaac 360
Db 460 TACCAATATGGAGACACCTATACCCAGGTTTGGGGCACCGAGATGTGGAACCCCAAC 519
Qy 361 cgtgagctgagcaggagctgctgtacctcaacgctgtggacaccatacccccggcctaca 420
Db 570 CGTGAGTGACGAGGAGTGTCTGTACCTCAACGTGTGGACACCATACCCCGCGCTTACA 579
Qy 421 tccccacccctgtctctgtgtgatctatgaggggtggtctctacagtggggacctctcc 480
Db 580 TCCCCACCCCTGTCTCTGTGTGATCTATGGGGGTGCTTCTACAGTGGGGCCTCTCTCC 639

Qy 481 ttgacgctgtacgatggcgcttcttgttacagccgagagagactgtctggtgtccatg 540
Db 640 TTGGACGTGTACGATGTCGCTTCTGTGACAGCCGAGAGACTGTCTGTGTGTCATG 699
Qy 541 aacLaccgggtgagagcctttggcttctctgcccgcgggagccgagagagagcccgagc 600
Db 700 AACATACCGGGTGGAGCCCTTGGCTTCTGTGCCCTTCCCGGGAGCCGAGAGCCCGGGC 759
Qy 601 aatgtgggtctctcttgatcacagagctggccctgcagtgagggtgcagggagaaagtgagcc 660
Db 760 AATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGCGAGCC 819
Qy 661 ttccgggggtgacccgacatcagtgacgctgtttggggagagcgaggagcgcgctcggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879
Qy 721 ggcatgacacctgtctcccccagccgagccggtcttccacagggcggtgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCGCCACAGCCGGGCGCTTTCACAGGGCGCTGCTGCAGAGC 939
Qy 781 ggtgcccccaatgagacctggccacctggggcagtggggcatgggagagggccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTGGGCCACGCTGGCATGGGAGAGGCCCTCGCAGGGCCACG 999
Qy 841 cagctggccaccctgtgggtgtctctccagggcgacgtgtgggaaatgacacagagctg 900
Db 1000 CAGTGGGCCCACTTGTGGGTGTCTCCAGCGGCGACTGTGGGAAATGACACAGAGCTG 1059
Qy 901 gtgacctgctctcgacacgacacgacgagcaggtcctggtgaaacacgaatgacacgtgctg 960
Db 1060 GTAGCTTCTCTTGGGACACAGCAGCGGAGGTCTCTGTGTGAACCACTAATGGCACCTGCTG 1119
Qy 961 cctcaagaaagcgtcttctccggttctctctgctgtggtagatggagagacttctcag 1020
Db 1120 CCTCAAGAAAGCGTCTCCGGTTCCTCTCTGTCCTGTGTGTGTGTGTGTGTGTGTGTGT 1179
Qy 1021 gacccccagagggccctcatcaagcggggagacctccagcgccctgcaggtgctggtggt 1080
Db 1180 GACACCCCGAGAGGGCCCTCATCAACGCGGGAGACTTTCACAGGCCCTGCAGAGTGTGTGTGGT 1239
Qy 1081 gtgggtgagaggtgagggctgtatttctgtttacgggggcccaggtctcagaaagac 1140
Db 1240 CTGCTGAAGGATGAGGGCTGCTATTCTGTGTTTACGGGGGCCCGAGGCTTACGAAAGAC 1299
Qy 1141 aacagagctctcatcagccgggcccagttcctggccgggtgctgggggttcccccag 1200
Db 1300 AACGAGTCTCTCATCAGCGGGCGGAGTTCCTGGCGGGGTGCGGGTTCGGGGTTCGCCAG 1359
Qy 1201 gtaagtgcacctggcagccgagagctgtgtcctgcaattacacagagactgctgcaccccgag 1260
Db 1360 GTAAAGTGAACCTGGCAGCGGAGGCTGTGTCTGTGCATTACACAGACTGTGATCCCGAG 1419
Qy 1261 gacccggcagccctgagggagggccctgagcagatgtgtggcgacaccaaagtctgtgc 1320
Db 1420 GACCCGGCACCCCTGAGGGAGGCCCTGAGCGATGTGTGTGGCGGACCAATGTCGTGTGTC 1479
Qy 1321 cccgtggccccagctggctggggcagctggctcccaggggtgcccgggttctacgctcagctc 1380
Db 1480 CCCGTGGGCCAGCTGGCTGGCGAGTGGCTGCCAGGGTGCCTGGGGTTCACGCCCTACGCTC 1539
Qy 1381 tttgaaacacctgcttccacgctctcctggccctgtggatgggggtgtcccacagggctac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTCTGGCCCTGTGGATGGGGGTGCCCGACGGCTAC 1599
Qy 1441 gagatcagagttcatctttgggatccccctggacccctctcgaaactcacacgagagagag 1500
Db 1600 CAGATCGAGTTCATCTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGAG 1659
Qy 1501 aaaattctgccccagcagctgatgcga tactgggggcaactttgcccgcacaggggagtcgcc 1560
Db 1660 AAAATCTTCGCCACGCGACTGATGCGATACTGGGGCAACTTTTGGCCCGCACAGGGGATCCC 1719
Qy 1561 aatgagccccgagagaccccccaagcccccaaatgtggccccgtatcacagggggggtcagcag 1620

Db	1720	ANTGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCGTACAGCGGGGGGCTCAGCAG	1779
Qy	1621	tacgttagcttggaacctgcggccgctgagagtgcggggggctgcgcgcagccagcctgc	1680
Db	1780	TACGTTAGTCTGACCTGCGCGCGCTGGAGTGCGCGGGGGCTGCGCGCCAGGCGCTGC	1839
Qy	1681	gcctctggaaccgcttctctcccaaaattgctcagcgcacc	1722
Db	1840	GCCTTCTGGAACCGCTTCTCCCAAAATTGCTCAGCGCCACC	1881
RESULT	5		
AR070207			
LOCUS		3016 bp	DNA
DEFINITION		Sequence 7 from patent US 5891725.	linear
ACCESSION		AR070207	
VERSION		AR070207.1	GI:7221095
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		1 (bases 1 to 3016)	
AUTHORS		Soreq,H., Zakut,H. and Eckstein,F.	
TITLE		Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them	
JOURNAL		Patent: US 5891725-A 7 06-APR-1999;	
FEATURES		Location/Qualifiers	
source		1..3016	
BASE		497 a	1065 c
ORIGIN		840 g	614 t
Query Match		99.88;	Score 1722; DB 6; Length 3016;
Best Local Similarity		100.0%;	Pred. NO. 3.2e-294;
Matches 1722;		Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	atgaggccccgcagtgctgtgcacacgcttccctggcttcccaactcttctctc	60
Db	160	ATGAGGCCCGCAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCACTCCTTCTCCTC	219
Qy	61	ctcctctggctctctgggtgagagagtgagggtcagggccggagagtcagagctgctg	120
Db	220	CTCCTCTGGCTCCTTGGTGAGGAGTGGGGCTGAGGGCGGGAGATGACAGGCTGCTG	279
Qy	121	gtgacgtgctggggccgcgctgcggggcattgcgctgaagacccccggggccctctc	180
Db	280	GTGACGGTGTGGGGGGCGGCTCGGGGGCAATCGCCTGAAGACCCCGGGGGCCTGTCTC	339
Qy	181	tctgtcttcttggcatccctcttgcgagaccacctagggaccctgcgtcttctgcga	240
Db	340	TCTGCTTTCCGTGGGCAATCCCTTTGGGAGGACCCCATGGGACCCCGTCGGTTTCTGGCA	399
Qy	241	ccggagcccaagcagccttggtcaggggtgtagacgtacaaaccttcagagtgctgc	300
Db	400	CCGGAGCCCAAGAGCCCTTGTCTGAGGGGTGTAGACGCTACAACCTTCCAGAGTCTCTGC	459
Qy	301	taccaaatgtggaacacctataccagggtttttaggggcaaccgagatgtgaaccccaac	360
Db	460	TACCAATATGTGGACACCCCTATACCAAGTTTTGGGGCACCCGAGATGTGGAACCCCAAC	519
Qy	361	cgtgagctgagcaggactgcttacctactcaactgtgtgacacatacccccgccctaca	420
Db	520	CGTGACCTGACCGAGGACTGCTGTACCTCAACGTGTGGACCACTATACCCCGGCTTACA	579
Qy	421	tcccccaacctctctcgtctggtatctatgagggtggtcttctacagtggggcctctccc	480
Db	580	TCCCCACCCCTGTCTCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCCTCTCTCC	639
Qy	481	ttgacgtgtacgatggccgcttcttggtaagcccgagagactgtgtggttccatg	540
Db	640	TTGACGCTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGACTGTCTGTGTTCATG	699

Qy 1621 tacattagctgagacctgcggccgctgaggtcgggcgggggtcgcgcccgagcctgc 1680
|||||
Db 1780 TACGTTAGCTGGACCTGCGGCCCTGAGAGTGGCGGGGCTGCGCGCCAGGCCTGC 1839
|||||
Qy 1681 gcctcttggaaccgtctctcccaaaattgctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAATTGCTCAGCGCCACC 1881
|||||
RESULT 6
AX275254
LOCUS AX275254 5767 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 3 from Patent WO0171014.
ACCESSION AX275254
VERSION AX275254.1 GI:16547674
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE
1 (sites)
AUTHORS Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic
plants
JOURNAL Patent: WO 0171014-A 3 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)
FEATURES
source Location/Qualifiers
1..5767
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM034"
BASE COUNT 1390 a 1521 c 1495 g 1361 t
ORIGIN

Query Match 99.68; Score 1718.6; DB 6; Length 5767;
Best Local Similarity 99.88; Pred. No. 1..ie-293;
Matches 172; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 atgaggecccgagtgctgctgcacacgcttccctggcttcccactctctctc 60
|||
Db 838 ATGGCTCCCGCAGTGCTGCTGCACACGCGTTCCTGGCTTCCCACTCTCTCCTC 897
|||||
Qy 61 ctctcttgcttcttggttgaggagtgaggctgagggcgggagatgcagactgctg 120
|||||
Db 898 CTCTCTGCTGCTTGGGTGGAGAGTGGGGCTGAGGGCGGGAGGATGCAGAGCTGCTG 957
|||||
Qy 121 gtgacggtgctggggcgctgctggggcattcgcttgaagaccccggggcccctgtc 180
|||||
Db 958 GTGACGCTGCTGGGGCGGCTGCGGGGCAATTCGCTGAAGACCCCGGGGGCCCTGTC 1017
|||||
Qy 181 tctgcttctctgggcatccctcttgcggagccaccatgggaccccgctgcttctgcga 240
|||||
Db 1018 TCTGCTTTCTTGGGCATCCCTTTGGGAGCACCCATGGGACCCGTCGCTTTCTGCCA 1077
|||||
Qy 241 ccggagcccaagacgcttgctcagggttgtagcgtacacgttccagagtgctgc 300
|||||
Db 1078 CCGAGGCCAAGACGCTTGTGTGAGGGGTGTGACGCTTACAACCTTCCAGAGTGTCTGC 1137
|||||
Qy 301 taccaatatgtggacaccctataccaggttttgagggcaccgagatgtgaaccccaac 360
|||||
Db 1138 TACCANAATGTGGACACCTATATCCAGGTTTGTAGGGCACCCGAGATGTGAACCCCAAC 1197
|||||
Qy 361 cgtgagctgagcaggagctgctgtacctcaacgctgtgagaccataccccggcctaca 420
|||||
Db 1198 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGACACCATACCCCGGCTTACA 1257
|||||
Qy 421 tccccaccctctgctctgttgatctatgagggtggtcttctacagtggggcctctcc 480
|||||
Db 1258 TCCCCACCCCTGCTCTGCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCCTC 1317
|||||
Qy 481 ttggacgtgtacgattggccgcttcttgtagagcccgagagactgtgctggttccatg 540
|||||

Db 1318 TTGGACGTTGACGATGGCCGCTTCTTGGTACAGCCGAGAGACTGTGCTGGTGTCCATG 1377
|||||
Qy 541 aactacccgggtgggagccttttggcttcttccctgcccgggagccgagagagcccgggc 600
|||||
Db 1378 AACTACCGGGTGGGAGCCTTTGGCTTCTTGCCCTTGCGGGAGCCGAGAGCCCGGGC 1437
|||||
Qy 601 aatgtgggtctctctggatcagaggtgctggccctgcagtggtgtcagggagaaagtcggcagccc 660
|||||
Db 1438 AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAGACGTGTCAGCC 1497
|||||
Qy 661 ttccgggggtgaccgaacatcagtgctgttttggggagagcgggagcgcctcggtg 720
|||||
Db 1498 TTCGGGGGTGACCCGACATCAGTACGCTGTTTGGGGAGAGCGCGGGAGCGCCTCGGTG 1557
|||||
Qy 721 ggcattgcacctgctgtcccccagcggcggcctgttcccacagggcgctgtcgcagagc 780
|||||
Db 1558 GGCATGCACCTGCTGTCCCGCCGACGCGGGCCTGTTCCACAGGGCCGCTGCTGCAGAGC 1617
|||||
Qy 781 ggtcccccaatggaacctgggccaagtcgggcatgaggagggccgctcgcagggccacg 840
|||||
Db 1618 GGTGCCCCCAATGGACCTGGGCCACGCTGGGCATGGAGAGGCCGCTCGCAGGGCCACG 1677
|||||
Qy 841 cagctggccaccctgtggtgcttccctccagcgcgacctggtgggaatgacacagagctg 900
|||||
Db 1678 CAGCTGGGCCCACTTGTGGGCTGTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG 1737
|||||
Qy 901 gtacgctgccttcggacacgaccagcgaggtccttgggtgaaccacgaatggcacgtgctg 960
|||||
Db 1738 GTAGCTGCTCTTCGACACAGCACAGCGAGGTCTGTTGGTGAACACGAATGGCACGTGCTG 1797
|||||
Qy 961 cctcaagaaagcgtctctcgggttctctcgtgctgtcgtgtggtagatgaggagattctcagt 1020
|||||
Db 1798 CCTCAAGAAAGCGCTTCTCCGGTCTCTCTGCTGCTGCTGTTAGTAGTAGAGACTTCTCAGT 1857
|||||
Qy 1021 gacccccagagccctcatcaacgggagagacttccagggcctcagagtcgtggtggt 1080
|||||
Db 1858 GACACCCAGAGGCCCTCATCAACCGGGAGACTTTCACAGGCTTCAGAGGTGCTGTGTGGT 1917
|||||
Qy 1081 gtggtgaaggatgagggtcgtctattttctggttttacggggggcccgaggtcttcagaaagac 1140
|||||
Db 1918 GTGTGAAGGATGAGGGCTCGTATTCTTCTGTTTACGGGGCCCCCAGGCTTCAGCAAGAC 1977
|||||
Qy 1141 aacgagctctcactaagccgggagtcctcgtgcccgggtgtcgggttgcgggttccccag 1200
|||||
Db 1978 AACGAGTCTCTCATCAGCGGGCGGAGTTCTTGGCGGGGTGCGGGTTCGGGTTCCCCAG 2037
|||||
Qy 1201 gtaagtgcctggcagccgagagctgtgctcgtgattacacagactgctcattccccag 1260
|||||
Db 2038 GTAAGTGACCTGGCAGCCGAGGCTGTGCTCTGCAATTACACAGACTGGCTGCATCCCCGAG 2097
|||||
Qy 1261 gaccgggcacgctgaggaggccctgagcgatgtggtggcgaccccaaatgtcgtgtgc 1320
|||||
Db 2098 GACCGGCACGCTTGAGGAGGCCCTGAGCGATGTGGTGGCGACCACAATGTCTGTGTC 2157
|||||
Qy 1321 cccgtggcccgagctggtcggcgagctgctgcccaaggtgcccgggttctacgcctacgtc 1380
|||||
Db 2158 CCCGTGGCCCAAGCTGGCTGGCGGACTGGCTGCCAGGTGCCCGGGTCTACGCCCTACGCTC 2217
|||||
Qy 1381 ttgaacaccgtgcttccacgctctcctggccctgtgagtggggtgcccacaggtctac 1440
|||||
Db 2218 TTTGNAACACCGTGTCTCCACGCTCTCTTGGCCCCCTGTGGATGGGGGTGCCCCACCGGCTAC 2277
|||||
Qy 1441 gagatcgaggttcatcttttgggatccccctggacccctctcgaatacacgcgagagag 1500
|||||
Db 2278 GAGATCGAGTTTCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGCGACAGGAG 2337
|||||
Qy 1501 aaaaacctcgcgcagcagctgatgcga tactggggccaactttgcccgcacaggggagtcacc 1560
|||||
Db 2338 AAAATCTTCGCCACGACTGATGGGACTGTGGGCAACTTTTGGCCGACAGGGGATCCC 2397
|||||
Qy 1561 aatgagccccagaccccaaggcccccaaatggccccgttacacggcgggggtcagcag 1620
|||||

Db 2398 AATGACCCCGAGACCCCAAGCCCCACAAATGGCCCCCTACACGGCGGGGGCTCAGCAG 2457

Qy 1621 tacgttagtctgacactgagccgctggaggctggaggtgcgcgggggtgcgcgccagggcctgc 1680

Db 2458 TAGCTTAGTCTGACCTGCGCGCGCTGAGGTGCGCGGGGGCTGCGCGCCAGGCGCTGC 2517

Qy 1681 gaccttgaaacgctctctcccccataattgctcagcgccacctga 1725

Db 2518 GCCTTCTGGAACCGCTTCTCCCAATTTGCTCAGCGCTACCTGA 2562

RESULT 7

AX275255 LOCUS AX275255 14446 bp DNA linear PAT 29-OCT-2001

DEFINITION Sequence 4 from Patent WO0171014.

ACCESSION AX275255

VERSION AX275255.1 GI:16547675

KEYWORDS .

SOURCE synthetic construct.

ORGANISM artificial sequence.

REFERENCE 1 (sites)

AUTHORS Mor,T.; Soreq,H.; Arntzen,C. and Mason,H.

TITLE Expression of recombinant human acetylcholinesterase in transgenic plants

JOURNAL Patent: WO 0171014-A 4 27-SEP-2001;

BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,

Tsafrir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,

Hugh S. (US)

FEATURES

source Location/Qualifiers

1..14446

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="plasmid vector pTM036"

BASE COUNT 3231 a 3031 c 4046 g 3042 t 296 others

ORIGIN

Query Match 99.6%; Score 1718.6; DB 6; Length 14446;

Best Local Similarity 99.8%; Pred. No. 9.2e-294;

Matches 1721; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 atgagggccccgcagtgctgtgtgacacgcttccctggttcccccaactcttctcctc 60

Db 142 ATGGCTCCCGCCAGTGTGTGTGACACGCTTCCCTGGCTTCCCACTCCTCTCTCCTC 201

Qy 61 ctctcttgctctggtggtgagagtgggggtgagggcgagagatgcagagtgctgc 120

Db 202 CTCCTCTGGCTCTCTGGGTGAGAGTGGGGGTGAGGGCGGGAGGATGCAGAGCTGCTG 261

Qy 121 gtacaggtgctggtggggcggtgcggggcattcgctgaagacccccggggccctgtc 180

Db 262 GTGACGGTGGTGGGGCGGGCTGCGGGCATTCGGCTGAAGACCCCCGGGGCCCTGTC 321

Qy 181 tctgcttctctgggatacccttggggagcaaccatgggagccctgcgtcttctgcga 240

Db 322 TGTGCTTTCTGGGCATCCCTTTGCGAGCCACCCATGGAGCCCGCTGCTCTTCTGCCA 381

Qy 241 ccggagcccaagcagccttggttcagggtgtgtagagctacacacctccagagtgtctgc 300

Db 382 CCGGAGCCCAAGCAGCTTGGTCAGGGGTGTGAGAGCTACAACCTCCAGAGTGTCTGC 441

Qy 301 taccatatgtgacaccctataccacagattttgagggcaccagagatgtggaaaccccaac 360

Db 442 TACCAATATGTGACACCCCTATACCCAGGTTTTTGGGGCACCCCATGGAGATGTGGAACCCCAAC 501

Qy 361 cgtgagctgagagagactgcctgtacctcaacgtgtgacaccataccccgggcctaca 420

Db 502 CGTGAGCTGAGCGAGGACTGCTGTGTACCTCAACGCTGTGACACCATATACCCCGGCTTACA 561

Qy 421 tcccccaacccctctctctgtctgtgattctatgggggtgggttctctacagtggggctctctcc 480

Db 562 TCCCCCAACCCCTGTCTCTGCTGTGGATCTATGGGGTGGGCTTCTACAGTGGGGCTCTCTCC 621

Qy 481 ttggacgtgtacagatggccgctctcttgygtacagggcgagagactgtgtgtgtccatg 540

Db 622 TTGGACGTGTACGATGGCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 681

Qy 541 aactaacgggtgggagacctttggcttctctggccctgcgggggagcagagagccccgggc 600

Db 682 AACTACCGGGTGGGAGCCTTTGGCTTCTTGGCCCTCCCGGGAGCGAGAGCCCGCGGC 741

Qy 601 aatgtgggtctcctctgcatcagaggtggccctgcagtgagtgagtgaggaagcgtggcagcc 660

Db 742 AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAACGTGGCAGCC 801

Qy 661 ttccgggggtgaccccgacatcagtgacgctgtttggggagagcgcgggagcccgctcggtg 720

Db 802 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGAGCCGCTCGGTG 861

Qy 721 ggcatacctgctgtcccccgcagccgggggctcttccacagggccgtgtctgcagagc 780

Db 862 GGCATGACCTGCTGTCCCGGCCACGCGGGGCTGTTCACAGGGCCGTGTCTGCAGAGC 921

Qy 781 gdtgcccccaatggaccctggccacggtggccatgggagagagcccgctgcaggggccag 840

Db 922 GGTGCCCCCAATGAGCCCTTGGGCCAGGTGGGCATGGAGAGCCCTGCAGGGCCACG 981

Qy 841 cagctggccccacctgtgtgggtgtctctccagggcggtggtggggaattgacacagagctg 900

Db 982 CAGCTGGCCACCTTGTGGCTGTCTCCAGGGGCGACTTGGTGGGAATGACACAGAGCTG 1041

Qy 901 gtacgctgctctggacacgacgagcgaggtctctggtgaacacgaatggcagctgtgc 960

Db 1042 GTAGCCTGCTTTCGGACACGACGCGCAGGTCTCTGTAACACACGAATGGCAGTGTG 1101

Qy 961 cctcaagaagcgctctccggttctctctctgctgctgtgtagtagagagacttccctcagt 1020

Db 1102 CCTCAAGAAAGCGTCTTCCGGTCTCTTCTGTCCTTGGTGGTAGATGGAGACTTCTCAGT 1161

Qy 1021 gacacccagagggccctcatcaacggtggagacttccagggcgtgcaggtgctgtggtgggt 1080

Db 1162 GACACCCAGAGGGCCCTCATCAACGGGGAGACTTCCAGGGCTGCAGGTGTGTGGGT 1221

Qy 1081 gtggtgaagatgaggtcgtcgtattttctggttttaagggggccccaggtcttcagaaagac 1140

Db 1222 GTGGTGAAGGATGAGGGCTCGTATTTCCTGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1281

Qy 1141 aacgagctctctcatcagccggcgaggttctctggtggcggtggtgcgggttccccag 1200

Db 1282 AACGAGTCTCTATCAGCCGGCGGCGAGTTCCTGGCCGGGGTGGGGTTCGGGTTCGCCAG 1341

Qy 1201 gtaagtgaacctggcagcgaggtgtgtgtctctgcattacacagactggtgcatccccag 1260

Db 1342 GTAAGTGACTGGCAGCGAGGCTGTGGTCTCTGCATTTACACAGACTGGCTGCATCCCGAG 1401

Qy 1261 gaccggcagcctgagggagggccctgagagatgtgtgtggggagacacaatgtcgtgtgc 1320

Db 1402 GACCCGGCAGGCTTGAGGAGGGCCCTGAGCATGTGGTGGGGGACCACAAATGTGTGTGC 1461

Qy 1321 cccgtggccccagctggtggtggcagctggctgcccaggtgtcccgggtctacgcctacgtc 1380

Db 1462 CCCGTGGCCACGCTGGCTGGGGGACTGGCTGCCAGGGTGCCCGGGTCTACGCCCTACGTC 1521

Qy 1381 ttgtgaacaacgtgtcttccacgctctcctgtgccccctgtggtgtgggtgtccccacggttac 1440

Db 1522 TTTGAACACCGTGTTCACGCTCTCTGGCCCTCTGTGGATGGGGGTGCCCCACGCGCTAC 1581

Qy 1441 ggaatcgatttcattcttgggattcccccttgaccctctcgaacctacacgacagagggag 1500

Db 1582 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCGAGAGGAG 1641

Qy 1501 aaaaattctgcgccagcgactgagtagactggtggccaaactttgcccgcacaggggattccc 1560

Db 1642 AAAATCTCGCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1701


```
QY 1231 ctgattacacagactggtgtgcatcccgaggagaccggcgccgctgagggagggccctgagc 1290
|||||
Db 1141 CTGATTACACAGACTGGTGTGACCTAGGACCCGCGGCTGAGGACGCGCTGAGC 1200

QY 1291 gatgtggtgggacacacaaatgctgctgcccgtgcccagctgactggcgagctggt 1350
|||||
Db 1201 GACGTGGTGGGCGACACACAGCGTGTGCCCCGTGGCCAGCTGGCTGGGCGACTGGCC 1260

QY 1351 gccagggtgcccgggtctacgctacgtcttggaaacacgctgcttcccagctctctctg 1410
|||||
Db 1261 GCCAGGCTGCCGGGTCTATGCTAGCTCTTTGAACACCGGCTTCCAGCTCTCTCTG 1320

QY 1411 cccctgtgatatgggggtgcccacagctacgagatgagtttaatttcttgggataccccctg 1470
|||||
Db 1321 CGGCTGTGATGGGCGTGGCCACGCTACGAGTACGATTCATCTTCGGGCTGGCCTTG 1380

QY 1471 gacccctctgaaactacacgcgcagagagaaaattctgcgcagcagctgatgcgatac 1530
|||||
Db 1381 GAACCTCTGCTGAACCTACACCGAGGAGAGAGGATCTTTGCCAGAGATTTGATGAGATAC 1440

QY 1531 tgggccaactttgcccgcacacaggggatacccaatgagccccgagaccccccaagggcccaaa 1590
|||||
Db 1441 TGGGCCAACTTTGACGCACAGGGACCCGACGAGCCCCGCGACGCCAAGGCCCCGCGAG 1500

QY 1591 tggccccctacacgcgcgggggtcagcagtagctagctgtagctgagcctgcccgcgtggag 1650
|||||
Db 1501 TGGCCACCTACACGCGGGGGCGCAGCAGTACGTTAGCTGTGAACCTGGCGCGCTGGAG 1560

QY 1651 gtgcgccccggggtgcgcgccagcctgcgctcttgcgaacgctctcccccaaatg 1710
|||||
Db 1561 GTGCGGAGGGGCTGGCGCCCGCAGGCTGCGCTTCTGGAACCGCTTCTGCCCAAACTG 1620

QY 1711 ctacgcgccacc 1722
|||||
Db 1621 CTCAGCGCCACC 1632

RESULT 9
LOCUS AX306139 2089 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 890 from Patent WO0188188.
ACCESSION AX306139
VERSION AX306139.1 GI:17645432
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 890 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1..2089
/organism="Mus musculus"
BASE COUNT 409 a 638 c 596 g 446 t
ORIGIN

Query Match 74.3%; Score 1282; DB 6; Length 2089;
Best Local Similarity 84.0%; Pred. No. 1.4e-216;
Matches 1447; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1 atgagggcccccagtgctgtgtgcacagccttccctggtctccaccactcttctctc 60
|||||
Db 1 ATGAGGCCCTCCCTGGTATCCCTGCTATACACCTTCCCTGGCTTTCCACTCTCTTCCCTC 60

QY 61 ctctctggtctctggtgagagagtgggggtgagggcgagggagagatgcagagctgctg 120
|||||
Db 61 CTCTCTCTCTCTCTGAGGAGGGGCAAGGCTGTAGGGCGCGGGAAGACCCGCGAGCTGCTG 120
```

```
QY 121 gtgaacgggtgcgtggtggggccggctgctgctgaagaccacccccggggccctgtc 180
|||||
Db 121 GTGAGGGTTCGAGGGGCGCAGCTGAGGGGCATCCGCTGAAGGCCCTTGAGAGCCAGTC 180

QY 181 tctgcttctctgggcatcccccttggtagggagccacccatgggagccccgtctgctttctgcca 240
|||||
Db 181 TCAGCTTTTCTGGGCAATCCCTTTTGAGAGCCACCTGTGGGGCTCACGTAGATTATATGSCA 240

QY 241 ccggagcccaagcagccttggtaggggtggttagacgcttacaaccttccagagtgctgc 300
|||||
Db 241 CCAGAGCCCAAGCGGCCCTGGTCAGGAGTGTGGATGCTACACACTTCCAAATATGTCGC 300

QY 301 taccatatgtggacacccctataccccaggttttgaggccacccagagatgtggaaaccccaac 360
|||||
Db 301 TACACAGTACGTGGACACCCCTGTACCTGGGTTTGGGGTACTGAGATGTGAACCCCAAC 360

QY 361 cgtgagctgaagcagagactgctgttacctcaacgtgtgaacacataccccgggctaca 420
|||||
Db 361 CGAGAGTTGAGTGAAGACTGCCCTGTATCTTAATGTGTGGACACCATACCCAGACCTGCT 420

QY 421 tccccacacccctgtcctcgtctggtgattctatgggggtggcttctacagtggggctctccc 480
|||||
Db 421 TCTCCACACACTGTCTCTCATCTGATCTATGGGGTGGTTTCTACAGCGAGCGGCTCC 480

QY 481 ttgagcgtgtacagatgggcgctctcttggtagagccgagagagactgtctggtgtccatg 540
|||||
Db 481 TTGGATGTGTATGACGGCGGCTTTCCTGGCCAGGTTTGGGGAGCTGTGTGTGTTATCTATG 540

QY 541 aactaccgggtgggagccttggcttctcctggccctcagggagcggagagccccgggc 600
|||||
Db 541 AACTACCGAGTGGGAACCTTTTGGCTTCTTGGCCCTTACCAGGAAGCAGAGAAAGCCCTGGC 600

QY 601 aatgtgggtctcctggtatcagagctggccctgctgagtggggtgcagagagaacgtggcagcc 660
|||||
Db 601 AATGTAGGTCTGCTGGATCAACGGCTTGCTTGCATGGTGGCAAGAAATATGTCAGCC 660

QY 661 ttgggggtgtaacccgacatacagtagcgtcttggggagagcgggagccgctcgtggtg 720
|||||
Db 661 TTTGGGGGCGACCCGATGTAGTACTGTGTGGGGAGAGTGGGGTGCAGCTCCGCTG 720

QY 721 gcatcacctctgtcccccgccagccggccttctccacagggcctgctgcagagac 780
|||||
Db 721 GGCATGCACATACTGTCTCCCTGCCACAGAGGCTCTTCCACAGGCTGTCTCTCCAGAGT 780

QY 781 ggtgcccccaatggacccctggccacggtgggcatggggagagggccgctgcaggggccaag 840
|||||
Db 781 GGCACACCCCAATGGGGCCCTTGGGCCACTGTAGTGTCTGGAGAGGCCAGCGCAGGGCCACA 840

QY 841 cagctggcccaacctgtgggctgctcctcagcgcgactggtggggaatgacacagagctg 900
|||||
Db 841 CTGCTGGCCCGCCTTGTGGGCTGTCCCCAGGTGGCGCTGGTGGCAATGACACCCGAGCTG 900

QY 901 gtacgctgctctcgacacgacacgacgagctcctggtgaacacacgaatggcaagctgctg 960
|||||
Db 901 ATAGCCTGCTTGGAGACAAGGCCGCTCAGGACCTGCTGGACCCACAGTGGCACGTCCTG 960

QY 961 cctcaagaagcgtcttccgggttctcctcctgctgcttctggtgtagatggagacttccctcagt 1020
|||||
Db 961 CCTCAAGAAAGATCTTCCGATTTTCTCGCTGTGTAGAGGGGACTTCCCTCAGT 1020

QY 1021 gacacccagagagccctcatcaacgagggagagacttccacggcctgaggtgctggtgggt 1080
|||||
Db 1021 GACACACCGGAGGCTCTCATCAATACTAGTAGATTTTCAAGACCTGTGAGGTGTGGTGGGT 1080

QY 1081 gtggtgaagatgagggctctgatttctggtttacggggcccgagccttcagcaagac 1140
|||||
Db 1081 GTGGTGAAGACGAGGGCTCTCTACTTCTTGTGTACGGGTCCAGGCTTCCAGCAAGAC 1140

QY 1141 aacgagctctctcatcagccggcgaggttctctggtccgggggtgcggtctcccccag 1200
|||||
Db 1141 AATGAATCTCTCATCAGCGGGGCCAGTCTCTGGCTGGGTGGGATCGGTGTACCCCAA 1200

QY 1201 gtaagtgaactggcagccgagggctgtggttctctgcatatacacagagctggtgcatcccgag 1260
```



```
Db 721 GGCATGCACATAGTCTCCCTGCCACAGGAGGCTCTTCCACAGGCGTGTCTCTCCAGAGT 780
QY 781 ggTgcccccaatgacacctgggcccaggtggtgcatgggagagcccgctgcagggcccaag 840
Db 781 GGCACACCAATGGGCGCTGGGCGACCTGTGAGTGTGGAGAGCCAGGCGCAGGGCCACA 840
QY 841 cagctgagccacctgtggtgtctctcagagcgagcactggtgggaatgacacagagctg 900
Db 841 CTGCTGGCGCGCTGTGGGCTGTCCCCAGGTGGCGCTGGTGGCAATGACACCCAGCTG 900
QY 901 gtacgtgctcttcgacacagcagcgaggtctctggtgaacacacaaagggcagctgctg 960
Db 901 ATAGCTCTGTGAGGACAGAGCCGCTCAGACCTGTGTGACACAGTGGCAGCTCTCTG 960
QY 961 cctcaagaagcgcttccgggtctctctgctgctgctgctgctgctgctgctcagct 1020
Db 961 CCTCAAGAAATGATCTCCGATTTCTTCTGCTGCTGTGTGAGACGGGACTTCTCTCAGT 1020
QY 1021 gacacccagagggccctcatcaacggcgagagacttccagcgctgaggtgctggtgggt 1080
Db 1021 GACACACCGGAGGCTCTCATATATAGTGAGATTTTCAAGACCTGCAGGTGTGTGGGT 1080
QY 1081 gtggtgaagatgagggctcgattctctggtttacggggcccgagccttcagcaaaagac 1140
Db 1081 GTGCTGAAGGACGAGGCTCTCTACTTCTGTTTACGGGCTCCAGGCTTCAGCAAGAC 1140
QY 1141 aacgagctctctcatcagcgggcgaggtctctgctgctgctgctgctgctgctgctcag 1200
Db 1141 AATGAATCTCTCATACGCGGGCCAGTCTCTGGCTGGGTGGGATCGGTGTACCCCAA 1200
QY 1201 gtaagtgaactggcagcgaggtgctggtctctgctcatcaacagactggtgctgctcag 1260
Db 1201 GCAAGTGACTTGGCGCGGAGGCTGTGGTCTCTCATATACAGACAGTGTGTGCACCCCTG 1260
QY 1261 gacccggcagcctcagggagggccctcagcgatggtggtgggacacacaaatgctgtgc 1320
Db 1261 GACCTCTACTACCTGAGAGATGCCATGAGTGCAGTGTGTGGCGACACCAACAGTGTGTG 1320
QY 1321 cccgtggccagctgctggtggcgactggtgcccaggtgcccaggtgctcagcctacgtc 1380
Db 1321 CCTGTGGCCAGCTGTGGCGACTGTGGTGTGCCAAGGGGCGCGGCTCTATGCTACATC 1380
QY 1381 ttggaacacgtgctccagctctctcagcctgagcctgaggtgggtgcccagcgctac 1440
Db 1381 TTGAACACCGTCCCTCCACACTGACTTGGCCCTCTGTGATGGGTGGCCCTATGCGCTAT 1440
QY 1441 gagatgagttcatctttgggaTcccccctggacccctctcgaactacacggcagagag 1500
Db 1441 GAAATCGAGTTTCATTTGGGCTCCCTCCCTGGATCCCTCGCTGAACATACACACGGAGG 1500
QY 1501 aaaaatttcgcccagcgaactgatcgatgactgggccaactttgcccgcacaggggatccc 1560
Db 1501 AGGATCTTTGCTCAGGCACTTATGAATACTATGAACCAATTTTCCGCGCACAGGGGACCC 1560
QY 1561 aatgagcccgagaccccaagcccacaatggcccccgctacacggcggggctcagcag 1620
Db 1561 AATGACCCCTCGAGACTTCAAAATCTCACAGTGGCCACCGTACACCACTGCGCGCGACAA 1620
QY 1621 tacgttagtctggaactgcgccgctggaggtgctgctgaggggctgctgctgagcctgc 1680
Db 1621 TATGTGAGCTGAACCTGAAGCCCTTAGAGGTGCGCGGGGACTGCGCGCCAGACCTGC 1680
QY 1681 gcttctggaacggcttctctcccaaatgtctcagcgccacc 1722
Db 1681 GCCTTCTGGAATGCTTCTCCCAAAATGCTCAGCGCCACC 1722

RESULT 11
S50879 LOCUS S50879 2066 bp mRNA linear ROD 08-MAY-1993
DEFINITION acetylcholinesterase T subunit [rats, mRNA Partial, 2066 nt].
ACCESSION S50879
VERSION S50879.1 GI:262092
```

```
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2066)
AUTHORS Legay,C., Bon,S., Vernier,P., Coussen,F. and Massoulié,J.
TITLE Cloning and expression of a rat acetylcholinesterase subunit:
generation of multiple molecular forms and complementarity with a
Torpedo collagenic subunit
J. Neurochem. 60 (1), 337-346 (1993)
J3107932
MEDLINE Genbank staff at the National Library of Medicine created this
REMARK entry [NCBI gibbsq 121005] from the original journal article.
This sequence comes from Fig. 1.
FEATURES
Location/Qualifiers
1..2066
/organism="Rattus sp."
/db_xref="taxon:10118"
1..2066
/gene="acetylcholinesterase T subunit, AChE"
1..1845
/gene="acetylcholinesterase T subunit, AChE"
/feature="this sequence comes from Fig. 1; AChE"
/codon_start=1
/product="acetylcholinesterase T subunit"
/protein_id="AAB24586.1"
/db_xref="GI:262093"
/translaton="MRPHYPLHTPSLASPLLLSLGGCARAREGREDPOLLYVRV
GQALGIRLKAFGSPVSARLGLIPFAEPVGSRRFPPEPKRPSGLDITATQNVCIY
YQDLYPGEGTEEMNPNRELSDELCLVNMWTPYPRPTSPVLVIYGGYSGASS
LDVYDRELAQVEGTVLVSMTYRVGTGFLALPGREAPNGVLLDQRLAQWQENI
AAFGDPMSVTLFGESAGAAVGMHILSLPSRSLFRAVLQSGTNGPWATVSAGAR
RRATLARLVGCPGCGAGNDTELISCLTRPAQDLVDHEWHPQESLFRSFPVV
DGDFLSDTPDALINTCDFQDLQVGVKDEGSFLVGVPGFSKDNESLISRAQFLA
GVRIGVPAQASDLAAEAVALHYTDWLHPEDPAHLRDAMSADVGDHNVCPVAQLAGLA
AGARVATIEFRAHSTLFWLWMPVPHGTEIEFIFGLPLSLNTVVERIFPAQRLM
QWTVPARTGDPNDPRDSRPMPTTAAQYVSLNKLPLEVRRLGRLAQTCAFWNR
LPKLLSATDLDIAERQWKAEFHRSSVMVHWKNOFDHYSKQERGSDL"
BASE COUNT 394 a 629 c 590 g 453 t
ORIGIN
Query Match 74.0%; Score 1276.2; DB 10; Length 2066;
Best Local Similarity 83.8%; Pred. No. 1.4e-215;
Matches 1443; Conservative 0; Mismatches 278; Indels 0; Gaps 0;
QY 1 atgagccccccagtgctgtgcacacgccttccctggctcccccactcctctcctc 60
Db 1 ATGAGGCTCCCTGGGTATCCCTGCATACACCTCCCTGGCTTCTCCACTCTCTCCCTC 60
QY 61 ctctctggctctgggtgagagtggtgggctgagggcgagggagagatgcagagctgctg 120
Db 61 CTCTCTCCCTCTCTGGGAGGAGGGCAAGGCTGAGGGCGGGGAAGACCTCTCAGCTGCTG 120
QY 121 gtgacgggtgctggggcggtgctggggcattcgctgaagaccccccgggggccctgtc 180
Db 121 GTGAGGTTTCGAGGGGCGCAGCTGAGGGGATCCGCTGAAGGCCCTCTGAGGCCAGTC 180
QY 181 tctgtcttctgggacatcccccttgggagcaccatgggaccccgctgcttcttgcga 240
Db 181 TCAGCTTTCTGGGCATCCCTTTTCAGAGAGCCACCTGTGGGCTCAGCTAGATTATGCCA 240
QY 241 ccggagcccaagcagcctgtgtcaggggtgtgtagcgcctacaacaccttccagagtgctgc 300
Db 241 CCAGAGCCCAAGCGCCCTGGTCAGGAATATTGGATGCTACCACTTCCAAATATGCTGC 300
QY 301 taccatagttgacacccctatacccgaggttttggggcaccgagatgtgggaaccccaac 360
Db 301 TACCAATACGTGGACACCCCTGTACCCTGGTGTGGAGGTACCGAGATGTGGAACCCCAAT 360
QY 361 cgtgagctgagcagagcagctgctgtacctcaacagctgtggacaccatacccccggtcaca 420
```

Db	361	CGAGAGCTGAGTGAAGACGTCCTTTATCTTAATGTGTGGACACCATACCCACGCGCTACT	420
Qy	421	tcccccacccctgtctctgtctggatctatgggggtggcttctacagtgggggctctctcc	480
Db	421	TCCTCCACACCTGTCCTCATCTGGATCTATGGGGGTGGTTCTACAGTGAGGATCCTCC	480
Qy	481	ttggacgtgtacgatggccgcctcttttggtaaggccgagaggactgtctgggtgtccatg	540
Db	481	TTGGAGCTGTATGACGGCCGTTTCTCTGGCCAGGTTGAGGGAACCGTGTGGTATCTATG	540
Qy	541	aactacggggtgggagcctttggcttctctggccctgcgcggggagccgagaggcccccgggc	600
Db	541	AACCTACCGAGTGGGAACCTTTGGCTCTCTTGGCTCTACAGGAACGACAGAAGCCCTGGC	600
Qy	601	aatgtgggtctcttgatcacagagctggccctgcagtggtgtgacgaggaacgttggcagcc	660
Db	601	ANTGTAGGCTGCTGGATGCACGGCTTGCCTTGCNATGGGTACAAAGAAAATATCGCAGCC	660
Qy	661	ttcgggggtgacccgacatacagtgacgtctgtttgggagagcgcggagccgcctcggtg	720
Db	661	TTTGGGGGAGACCCAAATGTCACTGACTCTGTTTGGGAGAGTGACAGTGCAGCCTCAGTG	720
Qy	721	ggcatgcacctgctgtccccgcacagccgcgggacctgttccacagggcgctgctgcagagc	780
Db	721	GGCATGCACATCTGTCTCTGCCACAGCAGGACGCTCTTCCACAGGGCTGTCTCTGCAGAT	780
Qy	781	ggtgcgcccaatgtgacctgggccacggttggggcataggagagggccctgcagaggccacg	840
Db	781	GGCACACCNAATGGGCCCTTGGCCACTGTGAGTGGGGAGAGGCCAGCGCAGGGCCACA	840
Qy	841	cagctggccacacctgtgggctgtcctccaggcggcactggtgggaatgacacagagctg	900
Db	841	CTGCTGGCCCGCCTTGTGGGCTGTCCCCAGGTGGCGCTGTGTGGCAATGACACCGAGCTG	900
Qy	901	gtagcctgacctcggacacgacagcgaggtcctgtgtgaaccgcgaattggcacgctgtg	960
Db	901	ATATCTGTCTTGAAGAACAGGCCGCTCAGGACCTGTGTGGACACGAGTGGCATGTGCTG	960
Qy	961	ccctcaagaagcgtctccggttctcctctgtgcctgtgtgactgtgtgagacttccctcagt	1020
Db	961	CCTCAAGAAAGTATCTTCCGGTTTCTTCTGTGCCTGTGTGTGGACGGGGATTTCTTCAGT	1020
Qy	1021	gacacccacagaggccctcatcaacgcgggagacttccacggcctgcagggtgtcgtgggt	1080
Db	1021	GACACGCCGAGCCTCTCATCAATACTGSAGATTTTCAAGACCTGCAGGTGCTGTGGGT	1080
Qy	1081	gtggtgaaggatgagggtctgtatttctgttttacggggcccagagcttcaagcaagac	1140
Db	1081	GTGGTGAAGGACAGAGGGCTCTACTTTCTGTGTTTACGGGGTCCAGAGCTTCAGCAAGAC	1140
Qy	1141	aacgagctctcatcaagcggccagttctctgcgcgggtgcgggttcgggggtccccag	1200
Db	1141	ANTGAATCTCTCATACAGCCGGGCCAGTTCCTGCGTGGGTGGCGATCGGTGTACCCAA	1200
Qy	1201	gtaagtgaacctggcagccgagctgtgtctctgcaattacacagactgctgtcatccccag	1260
Db	1201	CGAGTGAACCTGCGGCCGAGGCTGTGTGCTTCGTATATACAGACTGCTGCTGACCTGAG	1260
Qy	1261	gacccggcacgcctgaggaggagccctgagcgaatgtggcggcgaccacaatgtcgtgtgc	1320
Db	1261	GACCTCTCCACCTGAGAGATGCCATGATGTGGGTGTGTAGCGCACCACAACAGTTGTGTGC	1320
Qy	1321	ccgttgcccagctgctgtgggcagctgactgcacagaggtcccgggttctacgctcactc	1380
Db	1321	CCTGTGCCCCAGCTGTGCTGGGGACTGTGCTGCCAAGGGGCTCGGGTCTATGCCCTACATC	1380
Qy	1381	tttgaacacgcgtcttccacgctctctctggccctctgtgaatgggggtgcgccacggctac	1440
Db	1381	TTTGAACACCGTGCCTCCACATTGACTTGGCCCTCTCGATGGGGGTGCCCATGGCTAT	1440
Qy	1441	gagatcagagttcatcttttggatcccccttgaccctctcgaaactacacggcagaggag	1500

Db	1441	GAATCGAGTTCACTCTTTGGGCTCCCTCGATGATCCCTCACTGAACACTACACCGTGGAGGAG	1500
Qy	1501	aaatcttcgcccagcagctagatgcgatactatggggccaaactttgcccgcacaggggatcccc	1560
Db	1501	AGATCTTTGCTCAGCAGCTATATGCACTACTGGACCAATTTTGGCCGCACAGGGGACCCC	1560
Qy	1561	aatgagccccgagaccocaaagcccccaaatggtcccccgctacacggtcggtcagcag	1620
Db	1561	AATGACCTCTCAGAGACTCTAAGTCTCCACGGTGGCCGCTACACCACTGCGCGCGCAGCAA	1620
Qy	1621	tacattagcttgaccctgcgcgcctagagtgctgcgcgggggctgcgcgccccagggctgc	1680
Db	1621	TACGTGAGCCCTGAACCTGAAGCCTTTGGAGTTCGGGGGGGACATGCGCGCCACACCTGC	1680
Qy	1681	gcctctggaaccgtctctctcccaaatgctcagcgcac	1721
Db	1681	GCCCTCTGGATCGTTTCTCTCCCAATGCTCAGCGCCAC	1721
RESULT 12			
HUMACHEB			
LOCUS	HUMACHEB	4185 bp	DNA linear PRI 23-JUN-1995
DEFINITION	Homo sapiens acetylcholinesterase (ACHE) gene, exons 2-6.		
ACCESSION	L42812		
VERSION	L42812.1	GI:854682	
KEYWORDS	acetylcholinesterase.		
SOURCE	Homo sapiens DNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (sites)		
TITLE	Bartels,C.F., Zelinski,T. and Lockridge,O. Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene accounts for yb blood group polymorphism		
JOURNAL	Am. J. Hum. Genet. 52 (5), 928-936 (1993)		
MEDLINE	93256075		
REFERENCE	2 (bases 1 to 4185)		
AUTHORS	Bartels,C.F., Moriarty,P.L., Becker,R.E., Robbs,R.S., Sorenson,R.C., Mountjoy,C.P. and Lockridge,O.		
TITLE	Polymorphic sites in the acetylcholinesterase gene of patients with Alzheimer's disease		
JOURNAL	Unpublished (1995)		
FEATURES	Location/Qualifiers		
source	1. .4185		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
exon	/map="7q22"		
	/cell_type="white blood cell"		
	75. .1162		
	/gene="ACHE"		
	/note="G00-118-746"		
gene	/number=2		
	/evidence=experimental		
	75. .4185		
variation	/gene="ACHE"		
	1151		
	/gene="ACHE"		
	/note="yt blood group; His322/Asn"		
intron	/replace="a"		
	1163. .1508		
	/gene="ACHE"		
	/note="G00-118-746"		
	/number=2		
	/evidence=experimental		
exon	1509. .1993		
	/gene="ACHE"		
	/note="G00-118-746"		
	/number=3		
	/evidence=experimental		
variation	1871		
	/gene="ACHE"		
	/note="Pro446/Pro"		
	/replace="t"		

```

Intron      1994..2988
            /gene="ACHE"
            /note="G00-118-746"
            /number=3
            /evidence=experimental
            variation
            /gene="ACHE"
            /note="g or a; within Alu repeat; g more common;
            polymorphism at 2309"
            /replace="a"
            2989..3158
            /gene="ACHE"
            /note="G00-118-746"
            /number=4
            /evidence=experimental
            Intron
            3159..3238
            /gene="ACHE"
            /note="G00-118-746"
            /number=4
            /evidence=experimental
            exon
            3239..3991
            /gene="ACHE"
            /note="Includes retained intron 5; G00-118-746"
            /number=5
            /evidence=experimental
            variation
            3290
            /gene="ACHE"
            /note="Pro 561 more common; C/G polymorphism at 3290;
            Pro561 (CG) or Arg 561 (CGG)"
            /replace="g"
            3992..>4185
            /gene="ACHE"
            /note="G00-118-746"
            /number=6
            /evidence=experimental
            BASE COUNT      792 a 1320 c 1236 g 837 t
            ORIGIN

Query Match      69.4%; Score 1198; DB 9; Length 4185;
Best Local Similarity 81.8%; Pred. No. 7.8e-202;
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

QY 1 atgagccccccgagtgctgtgtgacacgccttccctggcttcccaactctctctc 60
   |||||
DB 95 ATGAGCCCCCGAGTGTGTGTGACACAGCCCTTCCTGGCTTCCCACTTCCTCTC 154

QY 61 ctctctgctcttggtgagagtgaggctgagggccgagagacagagctgctg 120
   |||||
DB 155 CTCTCTGGCTCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGATGCAGAGCTGCTG 214

QY 121 gtgacgggtggtggggccggtcgggggcattcgctgaagacccccggggccctgtc 180
   |||||
DB 215 GTGACGGTGTGGGGCCGGCTGCGGGGCATTTCGGCTGAAGACCCCGGGGGCCCTGTC 274

QY 181 tctgcttctctgggacatcccttggagagcaaccatgggaccccgctcgcttctgca 240
   |||||
DB 275 TCTGCTTTCTTGGGATCCCTTTGCGGAGCCACCACCATGGACCCCGCTTCTGCA 334

QY 241 ccggagcccaagcagccttggtcaggggtgtagacgtacacaccttcagagtgctgc 300
   |||||
DB 335 CCGGAGCCCAAGCAGCTTGGTCAGGGGTGTAGAGCCTACACCTTCCAGAGTGTCTGC 394

QY 301 taccataatgtgacacccctataccacagtttttggaggcaccgagatgtggaaccccaac 360
   |||||
DB 395 TACCAATATGTGACACCCCTATACCCAGGTTTGTAGGGCACCAGATGTGGAACCCCAAC 454

QY 361 cgtgagctgagcagagactgcctgtacctcaacgtgtggacaccatacccccgccata 420
   |||||
DB 455 CGTGAAGCTGAGCGAGGACGTGCTGTACCTCAACGCTGTGACACCATACCCCGGCTACA 514

QY 421 tcccccaacctctctcgtctgagatctatgggggtgggtctctacagtggggctcctcc 480
   |||||
```



```
repeat_region 3799..4102 /rpt_family="Alu"
repeat_region 4550..4853 /rpt_family="Alu"
repeat_region 4879..5089 /rpt_family="Alu"
repeat_region 5107..5403 /rpt_family="Alu"
repeat_region 5495..5782 /rpt_family="Alu"
repeat_region 6357..6656 /rpt_family="Alu"
repeat_region 6661..6848 /rpt_family="Alu"
repeat_region 6849..7025 /rpt_family="Alu"
repeat_region 7770..8089 /rpt_family="Alu"
repeat_region 8110..8419 /rpt_family="Alu"
misc_feature 8185..8623 /note="similar to EST AI539956 (NID:g4457329)"
repeat_region 9602..9902 /rpt_family="Alu"
repeat_region 10324..10630 /rpt_family="Alu"
repeat_region 11217..11526 /rpt_family="Alu"
repeat_region 11667..11976 /rpt_family="Alu"
repeat_region 12357..12655 /rpt_family="Alu"
repeat_region 12657..12781 /rpt_family="Alu"
repeat_region 12783..12955 /rpt_family="Alu"
repeat_region 12956..13236 /rpt_family="Alu"
repeat_region 13237..13367 /rpt_family="Alu"
repeat_region 13403..13444 /rpt_family="Alu"
repeat_region 13445..13547 /rpt_family="Alu"
repeat_region 13561..13864 /rpt_family="Alu"
repeat_region 14215..14326 /rpt_family="Alu"
repeat_region 14327..14591 /rpt_family="Alu"
repeat_region 14592..14674 /rpt_family="L1"
repeat_region 14675..14966 /rpt_family="Alu"
repeat_region 14967..15001 /rpt_family="L1"
repeat_region 15010..15337 /rpt_family="Alu"
repeat_region 16240..16509 /rpt_family="Alu"
repeat_region 16522..16656 /rpt_family="Alu"
repeat_region 16687..16978 /rpt_family="Alu"
repeat_region 17053..17344 /rpt_family="Alu"
repeat_region 18167..18444 /rpt_family="Alu"
repeat_region 18848..19163 /rpt_family="Alu"
misc_feature 19101..19232 /note="similar to EST BE160273 (NID:g8622994)"
repeat_region 19172..19471

/rpt_family="Alu"
19482..19793 /rpt_family="Alu"
19889..20004 /rpt_family="Alu"
20155..20444 /rpt_family="Alu"
20455..20725 /rpt_family="Alu"

Query Match 69.4%; Score 1198; DB 9; Length 172358;
Best Local Similarity 81.8%; Pred. No. 3.7e-202;
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

QY 1 atgagccccccagtgctgtgcacacgccttccctggttcccccactccttctcctc 60
Db 136364 ATGAGCCCCCGCAGTGTCTGTGCACACGCCCTTCCCTGGCTTCCCCTCTCTC 136305

QY 61 ctctctgctcctggtgagagagtgagggtgagggcgagggatgcagagctgctg 120
Db 136304 CTCTCTGGCTCTGGGTGAGGAGTGGGGCTGAGGGCGGAGGATGCAGAGCTCTG 136245

QY 121 gtgacggtcggtggggccggtcggtggtggtggtggtggtggtggtggtggt 180
Db 136244 GTGACGGTGGTGGGGCGGCTGCGGGGCAATTCGGCTGAAGACCCCGGGGGCTGT 136185

QY 181 tctgcttctcctgggcatccccctttgcggagccaccatgggaccccgctgcttctgcca 240
Db 136184 TCTGCTTCTCTGGGCATCCCTTTTCGGGAGCCACCCATGGACCCCGTCTGCTTTCGCCA 136125

QY 241 ccggagcccaagcagccttgggtcaggggtggtgagcgtacacaccttccagagtgctgc 300
Db 136124 CCGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGAGCTACAACTTCCAGAGTGTCTG 136065

QY 301 taccatatgtgacacccctataccccaggttttgaggggcaccagagatggaaccccaac 360
Db 136064 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGGACCCGAGATGTGACCCCAAC 136005

QY 361 cgtgagctgagcgagagactgctgtacctcaacgtgtgacaccatacccccgccctaca 420
Db 136004 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTTGGACACATACCCCGGCTTACA 135945

QY 421 tcccccaacccctgtctctgttgatctatgggggtggtggttctacagtgagggtcctctcc 480
Db 135944 TCCCCACCCCTGTCTCTGTCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC 135885

QY 481 ttggacgtgtacgagtgccgcttcttgggtacagggcgagagagactgtgctgggtgctcatg 540
Db 135884 TTGGACGTGTACGATGGCCGCTTCTTGGTTACAGGCCGAGAGGACTGTGTGTGTCTCANTG 135825

QY 541 aactaccgggtgggagccttggccttctcctggcctccggggagcgagagcccgccgggc 600
Db 135824 AACTACCGGTGGGAGCCTTGGCTTCTTCCCTGGCCCTCCGGGAGCGAGAGCCCGCCGGGC 135765

QY 601 aatgtgggtctcctgagatcagaggtggtggtggtggtggtggtggtggtggtggtggtggt 660
Db 135764 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGACGTGGGTGGAGAGAACGTGGCAGCC 135705

QY 661 ttccgggggtgaccccgacatcagtgacgctgtttggggagagcgcgagggccgctcggtg 720
Db 135704 TTCGGGGGTGACCCCGACATCAGTGACGCTGTGTGGGGAGAGCGCGGGGCTCGGTG 135645

QY 721 ggcatacactgtcttccccccagcgggggctgttccacagggcctgtgctgcagagc 780
Db 135644 GGCATGCACTGTCTGTCCCGCCAGCCGCGGGGCTGTCTTCCACAGGCCCTGTCTGTGCAGAGC 135585

QY 781 ggtgcccccaatggaccctgggccacggtggtggtggtggtggtggtggtggtggtggtggt 840
Db 135584 GTTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCTGCCAGGGCCAGC 135525

QY 841 cagctggccacactgttgggtgctcctccagcgcgacactggtggggaatgacacagagctg 900
Db 135524 CAGCTGGCCCACTGTGTGGGCTGTCTTCCAGGGGCTGTCTTCCAGGGGCTGTGTGGGGAATGACACAGAGCTG 135465
```


* 19528	20570: contig of 1043 bp in length	
* 20571	20670: gap of unknown length	
* 20671	22193: contig of 1523 bp in length	
* 22194	22293: gap of unknown length	
* 22294	24473: contig of 2180 bp in length	
* 24474	24573: gap of unknown length	
* 24574	25995: contig of 1422 bp in length	
* 25996	26095: gap of unknown length	
* 26096	27843: contig of 1748 bp in length	
* 27844	27943: gap of unknown length	
* 27944	30120: contig of 2177 bp in length	
* 30121	30220: gap of unknown length	
* 30221	31961: contig of 1741 bp in length	
* 31962	32061: gap of unknown length	
* 32062	35719: contig of 3658 bp in length	
* 35720	35819: gap of unknown length	
* 35820	38071: contig of 2252 bp in length	
* 38072	38171: gap of unknown length	
* 41172	41171: contig of 3000 bp in length	
* 41272	47552: contig of 6281 bp in length	
* 47553	47652: gap of unknown length	
* 47653	52017: contig of 4365 bp in length	
* 52018	52117: gap of unknown length	
* 52118	55609: contig of 3492 bp in length	
* 55610	55709: gap of unknown length	
* 55710	60279: contig of 4570 bp in length	
* 60280	60379: gap of unknown length	
* 60380	66246: contig of 5867 bp in length	
* 66247	66346: gap of unknown length	
* 66347	78184: contig of 11838 bp in length	
* 78185	78284: gap of unknown length	
* 96185	96184: contig of 17900 bp in length	
* 96185	96284: gap of unknown length	
* 115729	115729: contig of 19445 bp in length	
* 115830	115829: gap of unknown length	
* 139050	139049: contig of 23220 bp in length	
* 139150	139149: gap of unknown length	
* 161903	161902: contig of 22753 bp in length	
* 161903	162002: gap of unknown length	
* 162003	194681: contig of 32679 bp in length.	
FEATURES		Location/Qualifiers
source		1..194681
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="7"
		/clone.lib="Caltech human BAC library C"
BASE COUNT		45257 a 49777 c 50799 g 45331 t 3517 others
ORIGIN		
Query Match		
Best Local Similarity 69.4%; Score 1198; DB 2; Length 194681;		
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;		
QY 1	atgagccccccagtgctgtgtgcacagccttccctggcttcccactcttctcctc 60	
DB 106067	ATGAGGCCCCCGCAGTGTCTGTGTGCACACGCTTCCCTGGCTCCCTCTCTCCTC 106126	
QY 61	ctcctctgctcctgggtgagagagtgagggtgagggcgagggcgagatgcagagctgctg 120	
DB 106127	CTCCTCTGGCTCTCTGGGTGGAGAGTGGGGGCTGAGGGCGGGAGATGCAGAGCTGCTG 106186	
QY 121	gtgagcgtgctggtggcgccgtcgaggcattcgctgaagacccccggggccctgtc 180	
DB 106187	GTGACGGTGTGGGGGGCGGCTGCGGGGCAATTCGGCTGAAGACCCCCGGGGCCCTGTG 106246	
QY 181	tctgcttctctgggcatccctcttgcggagccaccatggagcccccgctgcttctgccca 240	
DB 106247	TCTGCTTCTCTGGGCATCCCTTTGGGAGCCACCCATGGAGCCCGTCTGCTTCTGCCA 106306	
QY 241	ccggagcccaagcagccttggtcaggggtggtagacgctacaacctccagagtgctctgc 300	
DB 106307	COGGAGCCCAAGCAGCTTGGTCAGGGGTGGTAGAGCTACAACTTCAGAGTGTCTGC 106366	
QY 301	taccaatagtggacacccctataccccaggttttgaggccacgagagatgtggaaaccccaac 360	
DB 106367	TACCAATATGTGGACACCCCTATACCCAGGTTTTCAGGGCACCAGAGATGTGGAACCCCAAC 106426	
QY 361	cgtgagctgagcgaggaactgcctgtacctcaacgctgtggacacataccccgggacctaca 420	
DB 106427	CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCTTACA 106486	
QY 421	tcccccaacctgtcctgtctgtggtatctatgggggtggttcttcaagtggtgggacctctcc 480	
DB 106487	TCCCCCACCCCTGTCTCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC 106546	
QY 481	tggagcgtgacagatgcgccttcttggtacagcgacgagagactgctggtgtccatg 540	
DB 106547	TTGGAGCTGTACATGGCGGCTTCTTGGGTACAGCGCCGAGAGGACTGTGCTGGTGTCCATG 106606	
QY 541	aactaccgggtggagccttggccttccctggcctgcggggagcgagagagggcccgggc 600	
DB 106607	AACTACCGGGTGGAGGCTTTGGCTTCTTGGCCCTGCGGGGAGCCGAGAGGCCCGCGGC 106666	
QY 601	aatgtgggtctcctggtacagaggtgcctcctcagtggtgcaggagaaactggtgcagcc 660	
DB 106667	AATGTGGGTCTCTGTGATCAGAGGCTTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 106726	
QY 661	ttcgggggtgaccccgacatcagtgacgctgtttggggagagcgcgagcccgctcggtg 720	
DB 106727	TTTCGGGGGTGACCCGACATCAGTGACGCTGTGTTGGGGAGAGCGCGGGAGCCCGCTCGGTG 106786	
QY 721	ggcatgcacctgctgtccccgccacgcccgggctgttccacagggccgctgctgcagagc 780	
DB 106787	GGCATGCACCTGTGTCTCCCGCCACGCGGGGCTGTCTCCACAGGGCCGTGTGTGCAGAGC 106846	
QY 781	ggtgcccccaatggacccttggccacggtgggcaatgggagagcccgctgcaggggccaag 840	
DB 106847	GSTGCCCCCAATGGACCCTTGGCCACGCTGGGCATGGGAGAGCCGCTGCAGGGGCACG 106906	
QY 841	cagctggccacctgttggtgctgctccagcgccactggtggggaatgacacagagctg 900	
DB 106907	CAGCTGGCCCACTTGTGGGCTGTCTCCAGGGCGGCACTGGTGGGAATGACACAGAGCTG 106966	
QY 901	gtagcctgcttcggacacagcagcgaggtcctggtgaacacacgaatggcagctgctg 960	
DB 106967	GTAGCCTGCTTCGGACACAGCAGCGCAGGTCTGCTGTAACCAACCAATGGCAGCTGCTG 107026	
QY 961	cctcaagaagaagctcttcgggttctcctctgtgctgctgtggtgagatggagacttctcagt 1020	
DB 107027	CCTCAAGAAGAAGCGTCTTCCGGTTCTCTTCTGTCGCTGTGTGTAGATGGAGACTTCTCTCAGT 107086	
QY 1021	gacacccagagggcctctcatcaacgcgagagacttccagggcctg ----- 1065	
DB 107087	GACACCCAGAGGCCCTCTATCAACGGGGAGAGACTTCCAGGGCTGCAGGTAAGTACTAGTGGC 107146	
QY 1066	----- 1065	
DB 107147	TAGTGGCGTGAAGCTGGCTCTCTTGGTCCCAACGGTCCCTCCCTCCCTGCAGGGACC 107206	
QY 1066	----- 1065	
DB 107207	CAGGCATGAGGCTTCTCCAGGCCCATTTCCAGAACTCCCAAGATCTCTCCCTTGAGGGCT 107266	
QY 1066	----- 1065	
DB 107267	CAGATCCCCAGGGTGGTCAGCAGGGGACAGAGAGAAACCCACCATGGGTCTATTTTCTCTTT 107336	
QY 1066	----- 1065	
DB 107327	CTCTGCATCCCTCCCTGTATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 107386	
QY 1066	----- 1065	

D	b	107387	TCAATCTCTCGGCTTTTGTGTGTGCCAATCTGTTTCCTGTCACAGTACTGTCGTGCTGCCTGGT	107446
Q	y	1066	-----caggctggtgggttggtgaagatga	1094
D	b	107447	CGCTCCATCCACCACCCCTCTCCTCACCCCAGTGCTGGTGGTGGTGAAGGATGA	107506
Q	y	1095	gggtctgtattttctggtttacaggggcccaaggttcagcaaaacaacagagtctctcat	1154
D	b	107507	GGGCTCGTATTCTTGTGTTTACGGGGCCCCAGGCTTCAGCAAAGAACACACGAGTCTCTCAT	107566
Q	y	1155	cagcgagcgaggttctctgccgcgggttcgaggttcgggttcccagtaactgacctggc	1214
D	b	107567	CAGCCGGGCGGAGTCTCTGGCCGGGGTGGGGGTTCGCCAGGTAAGTGAACCTGGC	107626
Q	y	1215	agcagaggtgtggtctgctgatcacagagctggtcgtcatccccagagaccgcgcacgct	1274
D	b	107627	AGCCGAGGCTGTGCTGTGCATTACACAGAGTGGCTGCATCCCGANGACCCGGCAGCGCT	107686
Q	y	1275	gagggagccctgagcagatgtggtggcgaccacaatgtctgtgcccgtyggccccagct	1334
D	b	107687	GAGGAGGCGCCCTGAGCGATGTGTGGCGGACCAAAATGTCGTGTGCCCGTGGCCAGCT	107746
Q	y	1335	ggctggcgagactgctaccagaggtgccgggtctcagccttacgtctttgaacacctgac	1394
D	b	107747	GGCTGGGCGAGTGGCTGCCAGGTTGCCGGGTCTACGCCCTACTGCTTTGTAACACCCGTGC	107806
Q	y	1395	tccaacgtctctctggccctgtgtagtgggggtgccccaggtctacagatcgagttcat	1454
D	b	107807	TTCCACGCTCTCTTGCGCCCTGTGTGATGGGGTGGCCACCGCTACGAGATCGAGTTTCAT	107866
Q	y	1455	cttgggataccccctggacccctctgaactcacgacacgagagagaatacttcgcacca	1514
D	b	107867	CTTTGGGATCCCCCTGGACCCCTCTCGAAATACACGCGAGAGAGAAAATCTTCGCCCA	107926
Q	y	1515	gcgactatgcgatactggcccaactttgcccgcacaggg	1554
D	b	107927	CGGACTGATGCGATCTGGGCCAACTTTGCCCGCACAGGG	107966
R	E	SULT 15		
H	S	AFA002993/c		
L	O	CUS	34921 bp	DNA linear PRI 12-JUN-2000
D	E	FINITION	Homo sapiens cosmid from 7q22,	complete sequence.
A	C	SSION	AF002993	
V	E	RION	AF002993.1	GI:2735699
K	E	YWORDS	HTG.	
S	O	URCE	human.	
O	R	GANISM	Homo sapiens	
R	E	ERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
A	U	THORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
T	I	TLE	1 (bases 1 to 34921)	
J	O	URNAL	Shapira,M., Tur-Kaspa,I., Bosgraaf,L., Livni,N., Grant,A.D.,	
M	E	DLINE	Grissaru,D., Korner,M., Ebstein,R.P. and Soreq,H.	
P	B	UMED	A transcription-activating polymorphism in the ACHÉ promoter	
R	E	ERENCE	associated with acute sensitivity to anti-acetylcholinesterases	
A	U	THORS	Hum. Mol. Genet. 9 (9), 1273-1281 (2000)	
T	I	TLE	Genomic sequencing in 7q22 revealed a novel arsenite resistance gene	
J	O	URNAL	Unpublished	
R	E	ERENCE	3 (bases 1 to 34921)	
A	U	THORS	Drescher,B., Rosenthal,A., Hinzmann,B. and Korner,M.	
T	I	TLE	Direct Submission	
J	O	URNAL	Submitted (07-MAY-1997) Genome Analysis, Institute of Molecular	
F	E	A	Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany	
E	A	TURES	Location/Qualifiers	
S	O	URCE	1. .34921	
			/organism="Homo sapiens"	
			/db_xref="taxon:9606"	
			/chromosome="7"	

```

/existence-not_experimental
3781. .3934
/notes="M2EF, score = 97.3%"
/existence-not_experimental
3781. .3934
/notes="Xpound exon prediction, score = 99% (0%)"
/existence-not_experimental
3781. .3927
/notes="Alignment: CDS, |CG415001|85.7|112|258"
/existence-not_experimental
3812. .3960
/notes="BLASTN (EMBL), 110. .258 of emb|U41500|CG41500
Cricetulus griseus arsenite-resistance protein (ars2)
mRNA,"
/existence-not_experimental
4054. .4281
/notes="Region: Data base match"
/existence-not_experimental
4054. .4247
/notes="M2EF, score = 84.8%"
/existence-not_experimental
4191. .4220
/notes="Xpound exon prediction, score = 98% (0%)"
/existence-not_experimental
4191. .4220
/notes="BLASTX (GENPEPT), 58. .87 of gb|U41500|CGU41500_1
Cricetulus griseus arsenite-resistance protein (ars2)"
/existence-not_experimental
4248. .4436
/notes="Alignment: CDS, | Intron"
/existence-not_experimental
4437. .4593
/notes="M2EF, score = 97%"
/existence-not_experimental
4437. .4593
/notes="Xpound exon prediction, score = 89% (0%)"
/existence-not_experimental
4437. .4593
/notes="Alignment: CDS, |CG415001|86.6|457|613"
/existence-not_experimental
4466. .4626
/notes="BLASTN (EMBL), 453. .613 of emb|U41500|CG41500
Cricetulus griseus arsenite-resistance protein (ars2)
mRNA,"
/notes="Region: Data base match"
/existence-not_experimental
4467. .4519
/notes="BLASTX (GENPEPT), 87. .139 of gb|U41500|CGU41500_1
Cricetulus griseus arsenite-resistance protein (ars2)"
/existence-not_experimental
4594. .4788
/notes="Alignment: EST, | Intron"
/existence-not_experimental
4809. .4840
/notes="BLASTX (GENPEPT), 140. .171 of gb|U41500|CGU41500_1
Cricetulus griseus arsenite-resistance protein (ars2)"
/existence-not_experimental
4879. .4991
/notes="Alignment: CDS, | Intron"
/existence-not_experimental
4992. .5118
/notes="Alignment: CDS,
|CG415001|88.3|717|844|CG415002|88.6|1144"
/existence-not_experimental
5045. .5152
/notes="BLASTN (EMBL), 738. .845 of emb|U41500|CG41500
Cricetulus griseus arsenite-resistance protein (ars2)
mRNA,"
/notes="Region: Data base match"
/existence-not_experimental
5119. .5208
/notes="Alignment: CDS, | Intron"
/existence-not_experimental
5209. .5264
/notes="Alignment: CDS, |CG415001|89.3|845|900"

```

```

/misc_feature
/existence-not_experimental
complement(5238. .5431)
/notes="BLASTN (EMBL), 1. .194 of emb|G30851|HS1140038 human
chromosome 7 STS sWS3840: single read."
/notes="Region: Data base match"
/existence-not_experimental
5757. .14176
/notes="GC score = 489.57 (8420bp)"
/notes="Region: GC content"
/existence-not_experimental
5844. .7208
/notes="CpG island score = 0.79, GC = 69.50%, CpGs = 126"
/notes="Region: CpG island"
/existence-not_experimental
6019. .6471
/notes="CpG core score = 8.90"
/notes="Region: CpG island"
/existence-not_experimental
6454. .6508
/notes="homology = 89.10%, score = 38, counts = 11"
/rpt_type=tandem
/rpt_unit=cctcc

Query Match 65.3%; Score 1126; DB 9; Length 34921;
Best Local Similarity 81.2%; Pred. No. 2.5e-189;
Matches 1505; Conservative 0; Mismatches 0; Indels 349; Gaps 3;

QY 1 atgagccccccagtgctgtgacacgccttccctccctgctccacccctctctc 60
|||||
DB 10809 ATGAGGCCCCCGAGTGTCTGTGCACACGCCCTTCCCTGGCTTCCACACCTCTCTCTC 10750

QY 61 ctctctgctctctgggtgagagtggtggggtgagggccgggagagtgagagtgctg 120
|||||
DB 10749 CTCCTCTGGCTCTGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGTCTG 10690

QY 121 gtacgggtggtggggccgggtgctgggggattcgcctgaagaccccccggggccctgctc 180
|||||
DB 10689 GTACGGGTGCTGGGGCCGGCTGCGGGGCAATTCGCTGAAGACCCCGGGGGCCCTGTC 10630

QY 181 tctgcttctctgggaccccttctggagggccacccatggagcccgctcttcttccca 240
|||||
DB 10629 TCTGCTTCTCTGGGATCCCTTTGGGAGCCACCCATGGGACCCCTCTCTTCTGCCA 10570

QY 241 ccggagcccaagcagccttggtaggggtgtagagcctacaacccctccagtgctctgc 300
|||||
DB 10569 CCGGAGCCCAAGCAGCTTGGTTCAGGGGTGTAGACGCTACAACTTCCAGAGTGTCTGC 10510

QY 301 taccatattgacacccctatcccgaggtttgagggccacccagatgtggaaccccaac 360
|||||
DB 10509 TACCAATATTGGACACCCCTATACCCAGGTTTGTAGGGGCACCGAGATGTGGAACCCCAAC 10450

QY 361 cgtgagctgagcagagactcctgtacctcaacgctgtggacacccatcccccggcctaca 420
|||||
DB 10449 CGTGAGCTGAGCAGGAGTCTGCTGACCTCAACGCTGTGGACACCATACCCCGGCTTACA 10390

QY 421 tccccacccctctctcgtctgtgatctatggtgggtggtctctacagtggtggcctctcc 480
|||||
DB 10389 TCCCCACCCCTCTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC 10330

QY 481 ttggagctgtacatggccgcttctgtgtacagccgagagagagactgtgctgggtccatg 540
|||||
DB 10329 TTGGACGCTGACGATGGCCCTCTTCTTGTACAGCCCGAGAGGACTGTGCTGGGTGCCATG 10270

QY 541 aactaccgggtggagagccttctgtctctctggtccctgcccgggagccgagagggccgggc 600
|||||
DB 10269 AACTACCGGGTGGAGGCTTTGGCTTCTGGCCCTGCGGGGAGCCGAGAGGCCCGGGC 10210

QY 601 aatgtgggtctctctgagatcagagcctggtccctgcagtggtgaggaagacgtggagcc 660
|||||
DB 10209 AATGTGGGTCTCTCTGGATCAGAGGCTGCGCTGTCAGTGGGTGTCAGGAGAACTGGCAGCC 10150

QY 661 ttccgggggtgacccgacatcagtgacgctgtttgggagagcgaggagcgccctcggtg 720
|||||

```

```
Db 10149 TTGGGGGTGACCGACATCAGTGACGCTGTTTGGGAGAGCGGGAGCGCGCTCGGTG 10090
Qy 721 ggcataccctgctgtcccccagccagcggggacctgttccacagggccgctgctgcagagc 780
Db 10089 GGCATGCACCTGCTGTCTCCCGCCAGCGGGGCTGTTTCCACAGGGCGGTGTCAGAG- 10029
Qy 781 ggtgcccccaatggaacctgggccaagtgggcaagtgggcatgggagagggcccgctcgagggccacg 840
Db 10030 -----CGAGGGCCACG 10019
Qy 841 cagctggcccaacctgtgtggtgtctctccagcgcgacctggtgggaaatgacacagagctg 900
Db 10018 CAGCTGGCCCACTTGTGGGTGTCTCTCCAGCGGCACTGTGTGGGAATGACACAGAGCTG 9959
Qy 901 gtacgctgcttcggacacgaccagcgaggtcctctggtgaaccacgaatgcaacgtgctg 960
Db 9958 GTAGCCTGCCCTTCGGACACGACACAGCGAGGTCTCTGGTGAACCAACGAATGGCACGTGCTG 9899
Qy 961 cctcaagaaagcgtcttcctcgtctctcctcgtcctctggtgtagatggagacttctcagt 1020
Db 9898 CCTCAAGAAAGGCTTCCGGTTCCTCTCTGCTGTGCTGTGGTAGATGGAGACTTCCTCAGT 9839
Qy 1021 gacacccagggccctcatcaacgcgggagacttccacggcctg----- 1065
Db 9838 GACACCCAGAGGCCCTCATCAACGCGGAGACTTCCACGGCCTCAGGTAACAGTGCC 9779
Qy 1066 ----- 1065
Db 9778 TAGCTGGCGTGAAGTGCTCTCTGTGGTCCCAACGGTCCCTCCCTCCTCCTGACGGACC 9719
Qy 1066 ----- 1065
Db 9718 CAGGATGAGGGCTTCTCAGGCCCATGTACAGGGGACAGAGAAAGCCACCATGGGT 9659
Qy 1066 ----- 1065
Db 9658 CTATTTTCTCTTCTCTCATCCCTCCCTGATCTCGTCTCTCTCTCTCTCTCTCTCG 9599
Qy 1066 ----- 1065
Db 9598 GGTCTGTAACTGTCATCTCTCTGCTCTTGTCTGTCCATCTGTTTCTGTACTTGTG 9539
Qy 1066 -----cagtgctgtgggtg 1081
Db 9538 TGCTGTGCTGTCTGCTCTCATCCACCCCTCTCTCCCTCACCCCGAGGTGCTGGTGGGTG 9479
Qy 1082 tggtaagatgagggctcgtatttctggtttacggggcccgagccttcagcaaaagaca 1141
Db 9478 TGGTCAAGGATGAGGGCTCGTATTTCTTGTGTTACGGGGCCCGAGGCTTCAGCAAAAGACA 9419
Qy 1142 acgagctctctcatcagccggccgaggttctctggtggcggtggtgcggggttccccagg 1201
Db 9418 ACGAGTCTCTCATCAGCGGGGCGGAGTCTCTGGCGGGGTGCGGGTCTCCCGAGG 9359
Qy 1202 taagtacactggacgaggtggtggtcctgcatcacagactggctgcatccccagg 1261
Db 9358 TAAGTGACCTTGGCACCGAGGCTGTGGTCTCTGCATTACACAGACTGGGTGTCATCCCGAGG 9299
Qy 1262 accgggacgctgagggagccctgagcgtatggtggtgggacacacaaatgctgtgccc 1321
Db 9298 ACCCGGACGCTGAGGGAGGGCCCTGAGCGGATGTGGTGGGCGACCAATGTCTGTGTC 9239
Qy 1322 ccgtggccagctggtgggagactggctgccc-agggtgcccgggtctacgctacgtc 1380
Db 9238 CCGTGGCCAGCTGGCTGGGCGACTGGGTGCCCNAGGGTGCCCGGGTCTACGCCCTACGTC 9179
Qy 1381 ttgaaacccgtgtctcacgctctcctggtgcccctgtggatgggggtgccccacaggtac 1440
Db 9178 TTTGAACACCGTGTCTCAGGCTCTCTGGGCCCTCTGGATGGGGGTGCCCGACGGCTAC 9119
Qy 1441 gagatcgagttcatcttgggatccccctggacccctctcgaaacctcacacgagagag 1500
Db 9118 GAGATCGAGTTCACTTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 9059
```

Qy 1501 aaaatcttcgcccagcgactgatcgatactaggccaaactttgcccgcacaggg 1554
Db 9058 AAAATCTTGGCCCGAGCGACTGATGCGTACTGGGCCAACTTTGCCCGCACAGGG 9005

Search completed: August 31, 2002, 18:18:14
Job time: 16547 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:25:58 ; Search time 824.08 Seconds
(without alignments)
3593.920 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgagcccccagtgctct.....aattgtcagcccaactga 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0
Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_Q32802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1725	100.0	1725	24	AAS17549 Synthetic human ac
2	1722	99.8	1845	24	Human cDNA encodin
3	1722	99.8	2256	16	Human acetylcholin
4	1718.6	99.6	5767	24	Plasmid pTM034. S
5	1718.6	99.6	14446	24	AAS17548
6	1684.8	97.7	2253	11	AAQ05998
7	1313.6	76.2	1800	11	AAQ05999
8	1282	74.3	2089	24	ABI99790
9	1198	69.4	12113	24	AAS17492

10	741.8	43.0	1698	20	AAX86613	cDNA encoding an a
11	741.8	43.0	2478	20	AAX86614	cDNA encoding an a
12	380	22.0	2416	21	AAZ49470	Human wild type Bu
13	380	22.0	2416	21	AAZ49471	Human Butyrylcholi
14	378.4	21.9	2400	14	AAQ42496	Full-length human
15	322.8	18.7	2445	7	AAQ60111	Sequence encoding
16	322.8	18.7	2445	14	AAQ48393	Full-length foetal
17	200.4	11.6	4377	23	ABL19443	Drosophila melanog
18	185.6	10.8	764	14	AAQ60110	Sequence encoding
19	185.6	10.8	764	14	AAQ42495	Foetal brain clone
20	142.6	8.3	1614	16	AAQ0569	Recombinant bile s
21	142.6	8.3	1614	17	AAQ7862	Human wild-type bi
22	142.6	8.3	2166	16	AAQ0579	Encodes bile salt-
23	142.6	8.3	2166	17	AAQ7861	cdna of human bile
24	142.6	8.3	3018	17	AAQ34787	Human milk bile sa
25	141	8.2	2428	13	AAQ20252	Bile-salt stimulat
26	141	8.2	3018	12	AAQ14186	Human milk bile-sa
27	139.4	8.1	2184	15	AAQ71874	Sequence encoding
28	139.4	8.1	2344	22	AAF31110	Cholesterol estera
29	139.4	8.1	2428	17	AAQ35250	Human bile salt-st
30	139.4	8.1	2428	20	AAZ20300	Human bile salt-st
31	139.4	8.1	2428	24	ABA04403	Human milk bile sa
32	139.4	8.1	2428	24	AAQ18538	cDNA encoding Bile
33	139.4	8.1	2487	21	AAQ63883	cDNA encoding a hu
34	137.4	8.0	2428	15	AAQ83221	Bile salt-stimulat
35	130.4	7.6	1908	12	AAQ11977	Bovine pancreatic
36	122.6	7.1	2520	22	AAH57518	Human pancreas cel
37	120.2	7.0	1680	22	AAF25258	Nucleotide sequenc
38	120.2	7.0	2202	23	AAQ74460	DNA encoding novel
39	120.2	7.0	2484	21	AAF21884	Human breast and o
40	118.2	6.9	2076	22	AAQ08430	Human secreted pro
41	118.2	6.9	27314	23	ABL19442	Drosophila melanog
42	117.4	6.8	2101	21	AAQ00680	Human Hydrolyase pr
43	117.4	6.8	3776	22	AAQ08405	Human secreted pro
44	117.4	6.8	3824	21	AAQ6341	cDNA encoding a no
45	117.4	6.8	3824	22	AAQ21514	Human cDNA sequenc

ALIGNMENTS

RESULT 1
AAS17549
ID AAS17549 standard; DNA; 1725 BP.
XX
AC AAS17549;
XX
DT 25-FEB-2002 (first entry)
XX
DE Synthetic human acetylcholinesterase gene.
XX
KW Human; acetylcholinesterase; AChE; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI: 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
XX Claim 11; Page 41-42; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of acetylcholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents synthetic human AChE gene optimised for
CC expression in plants.
XX
SQ Sequence 1725 BP; 273 A; 553 C; 571 G; 328 T; 0 other;

Query Match 100.0%; Score 1725; DB 24; Length 1725;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagccccccagtgctgtcgacacgcttcctcctggtccctccctccctccctccctcc 60
DB 1 atgagccccccagtgctgtcgacacgcttcctcctggtccctccctccctccctcc 60

QY 61 ctctctggtctcctggtgagagtggtgggctgagggcgaggagatgcagagctgctg 120
DB 61 ctctctggtctcctggtgagagtggtgggctgagggcgaggagatgcagagctgctg 120

QY 121 gtgacggtgcgtggggcggtcggtggggcattcgctgaagacccccggggccctgtc 180
DB 121 gtgacggtgcgtggggcggtcggtggggcattcgctgaagacccccggggccctgtc 180

QY 181 tctgcttctcctggtccctctcgagacccacacacacacacacacacacacacacac 240
DB 181 tctgcttctcctggtccctctcgagacccacacacacacacacacacacacacac 240

QY 241 ccggagcccaagcagccttggtcaggggtgtagagcctacaaccttcagagtgctgc 300
DB 241 ccggagcccaagcagccttggtcaggggtgtagagcctacaaccttcagagtgctgc 300

QY 301 taccataatgtgaacacctataccacaggttttgaggggacccagagatgtggaacccaac 360
DB 301 taccataatgtgaacacctataccacaggttttgaggggacccagagatgtggaacccaac 360

QY 361 cgtgagctgagcagagactgctgtacccacacgctgtgacacacacacacacacacac 420
DB 361 cgtgagctgagcagagactgctgtacccacacgctgtgacacacacacacacacac 420

QY 421 tccccacacccctgctcgtctgtgatatgtgggggtgggttctacagtggggctcctcc 480
DB 421 tccccacacccctgctcgtctgtgatatgtgggggtgggttctacagtggggctcctcc 480

QY 481 ttggaagctgtacagatggccgctcttcttggttacagccgagagagactgtgctgggtccatg 540
DB 481 ttggaagctgtacagatggccgctcttcttggttacagccgagagagactgtgctgggtccatg 540

QY 541 aactaccgggtggagaccttgcttctcctggtccctccggggagcagagagccccgggc 600
DB 541 aactaccgggtggagaccttgcttctcctggtccctccggggagcagagagccccgggc 600

QY 601 aatgtgggtctcctggtacagaggtggccctcagtggtgcagagagaacgtgggaagcc 660
DB 601 aatgtgggtctcctggtacagaggtggccctcagtggtgcagagagaacgtgggaagcc 660

QY 661 ttccgggggtgacccagacatcagtgacgctgtttggggagagcgggagcgcctcggtg 720
DB 661 ttccgggggtgacccagacatcagtgacgctgtttggggagagcgggagcgcctcggtg 720

QY 721 ggcagcacctgctgtcccccagccagccgggggctgtgtccacagggccgctgctgcagagc 780
DB 721 ggcagcacctgctgtcccccagccagccgggggctgtgtccacagggccgctgctgcagagc 780

QY 781 ggtgcccccaatggacctgggccaaggtggcgatggagagagcccgctgcagggccacg 840
DB 781 ggtgcccccaatggacctgggccaaggtggcgatggagagagcccgctgcagggccacg 840

QY 841 cagctggccccactgtggtgctgtctctccagcggaactgtgtgggaatgacacagagctg 900
DB 841 cagctggccccactgtggtgctgtctctccagcggaactgtgtgggaatgacacagagctg 900

QY 901 gtacgtgctctcgacacagcagcagcagctgctgtggtgaacacacaaatggcagctg 960
DB 901 gtacgtgctctcgacacagcagcagcagctgctgtggtgaacacacaaatggcagctg 960

QY 961 cctcaagaaagcgtctctccggttctcctgtgctgtggtgtagatggagacttccctcagt 1020
DB 961 cctcaagaaagcgtctctccggttctcctgtgctgtggtgtagatggagacttccctcagt 1020

QY 1021 gacacccccagagccctcatcaacgcgggagacttccacggtcgaggtgctgggtgggt 1080
DB 1021 gacacccccagagccctcatcaacgcgggagacttccacggtcgaggtgctgggtgggt 1080

QY 1081 gtggtgaagagatgaggtcgtatttctggtttacggggccccaggtctcagcaaaagac 1140
DB 1081 gtggtgaagagatgaggtcgtatttctggtttacggggccccaggtctcagcaaaagac 1140

QY 1141 aacgagctctctcatcagcggggcgaggttctctggtggcggtgcgggttcgggttccccag 1200
DB 1141 aacgagctctctcatcagcggggcgaggttctctggtggcggtgcgggttccccag 1200

QY 1201 gtaagtgaacctggcagccgaggtgtggtctcattacacagactggtgcgtcccgag 1260
DB 1201 gtaagtgaacctggcagccgaggtgtggtctcattacacagactggtgcgtcccgag 1260

QY 1261 gacccggcagcctgagggagggccctgagcgtgtggtggggcagcacacaatgtcgtgtgc 1320
DB 1261 gacccggcagcctgagggagggccctgagcgtgtggtggggcagcacacaatgtcgtgtgc 1320

QY 1321 ccggtggccccagctggtggcgactgggtcccaggggtcccgggttctacgcctcagtc 1380
DB 1321 ccggtggccccagctggtggcgactgggtcccaggggtcccgggttctacgcctcagtc 1380

QY 1381 ttggaacacagctgcttccacgctcctcctggtccctgtggtgggtgggtgcccccggtac 1440
DB 1381 ttggaacacagctgcttccacgctcctcctggtccctgtggtgggtgggtgcccccggtac 1440

QY 1441 gagatcagttctatcttgggatccccctggacccctctcgaaactacacagcgagagag 1500
DB 1441 gagatcagttctatcttgggatccccctggacccctctcgaaactacacagcgagagag 1500

QY 1501 aaaaatttggccccagcagctgtagcgtactgtgggcaacttgcgcgcacaggggattccc 1560
DB 1501 aaaaatttggccccagcagctgtagcgtactgtgggcaacttgcgcgcacaggggattccc 1560

QY 1561 aatgagccccgagaccccccaagcccccaaatggcccccggtacacggcggtggtcagcag 1620

```

|||||
Db 1561 aatgagcccgagaccaccccaagcccccaaatggccccgtacacgcgcgggggtcagcag 1620
Qy 1621 tacgttagctgacctgcgcgcctgaggtgcggcgggggtcgcgcgcacagcctgc 1680
Db 1621 tacgttagctgacctgcgcgcctgaggtgcggcgggggtcgcgcgcacagcctgc 1680
Qy 1681 gcctctggaaccgtctctctcccaaatgtctcagcgccacctga 1725
Db 1681 gcctctggaaccgtctctctcccaaatgtctcagcgccacctga 1725

RESULT 2
AAS17493
ID AAS17493 standard; cDNA; 1845 BP.
XX AC
XX AAS17493;
XX
XX 26-FEB-2002 (first entry)
XX
XX Human cDNA encoding acetylcholinesterase, ACHE.
XX
XX Human; ss; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;
XX genotyping; neurological disease; Parkinson's disease;
XX Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..1845
XX /*tag= a
XX /product= "ACHE"
XX /tag= b
XX /tag= c
XX /standard_name= "Single nucleotide polymorphism"
XX replace (101,A)
XX /*tag= c
XX /standard_name= "Single nucleotide polymorphism"
XX replace (183,C)
XX /*tag= d
XX /standard_name= "Single nucleotide polymorphism"
XX replace (838,G)
XX /*tag= e
XX /standard_name= "Single nucleotide polymorphism"
XX replace (873,T)
XX /*tag= f
XX /standard_name= "Single nucleotide polymorphism"
XX replace (1032,A)
XX /*tag= g
XX /standard_name= "Single nucleotide polymorphism"
XX replace (1057,A)
XX /*tag= h
XX /standard_name= "Single nucleotide polymorphism"
XX replace (1431,T)
XX /*tag= i
XX /standard_name= "Single nucleotide polymorphism"
XX
XX WO200179219-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US11853.
XX
XX 14-APR-2000; 2000US-197173P.
XX
XX (GENA-) GENA/ISSANCE PHARM INC.
XX (KAZE/) KAZEMI A.
XX
XX Bentivegna SC, Chew A, Choi JY, Koshy B;
XX WPI; 2002-055248/07.
XX P-PSDB; AAU11231.
XX

```

```

PT New polymorphic variants comprising acetylcholinesterase (ACHE)
PT isogene, useful in expressing ACHE protein for use in screening for
PT candidate drugs to treat diseases related to ACHE activity, e.g.
XX neurological diseases or cancer -
PS Claim 26; Fig 2; 79pp; English.
XX
XX The invention relates to a polynucleotide comprising a polymorphic
XX variant of an acetylcholinesterase (ACHE) gene or fragment, protein or
XX complement, the variant comprising an ACHE isogene defined by a haplotype
XX selected from haplotypes 1-20 listed in the specification. Also included
XX are methods for haplotyping and genotyping the ACHE gene of an
XX individual, a method for predicting a haplotype pair for the ACHE gene of
XX an individual, a method for identifying an association between a trait
XX and at least one haplotype or haplotype pair of ACHE gene, recombinant
XX nonhuman organisms transformed or transfected with the polynucleotide
XX where the organism expresses ACHE protein encoded by the first
XX nucleotide sequence or encoded by the polymorphic variant sequence,
XX an isolated antibody specific for and immunoreactive with ACHE,
XX a method of screening for drugs targeting the polypeptide contacting ACHE
XX polymorphic variant with a candidate agent and assaying for binding
XX activity, a computer system for storing and analysing polymorphism data
XX for ACHE gene and a genome anthology for ACHE gene which comprises ACHE
XX isogenes defined by haplotypes 1-20 given in the specification.
XX The polymorphisms are useful for studying the biological function of
XX ACHE as well as in identifying drugs targeting this protein for the
XX treatment of disorder related to its abnormal expression or function.
XX The polymorphic variants may also be used in screening for compounds
XX targeting ACHE to treat a specific condition or disease predicted to be
XX associated with ACHE activity e.g. neurological diseases (e.g.
XX Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
XX and tumours. The ACHE gene maps to human chromosome 7q22. The
XX present sequence is the coding sequence of the ACHE gene.
XX
XX Sequence 1845 BP; 302 A; 590 C; 608 G; 345 T; 0 other;

```

```

Query Match 99.8%; Score 1722; DB 24; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagggcccgagctgtctgcacacgctctccctggtctccactctctctc 60
Db 1 |||||
Qy 61 ctctctggtctggtgggagagtgagggtggtgagggcgccggagagtgacagctgctg 120
Db 61 |||||
Qy 121 gtgacggtgcgtggggcgcggtgcggggttcctgcctgaagaccccggtggtc 180
Db 121 |||||
Qy 181 tctgtcttctgggcatcccttctgcgagcaccatcggtgaccccggtcttctgcca 240
Db 181 |||||
Qy 241 ccggagcccaagcagccttggtaggggtggtagcgtctacacacctccagagtgctgc 300
Db 241 |||||
Qy 301 taccaatatgtggacacccctataccagggttttggggcaccgagatgtggaccccaac 360
Db 301 |||||
Qy 361 cgtgagctgagcagagactgctgtacctcaacgctgggaaccataccccggcgctaca 420
Db 361 |||||
Qy 421 tccccaccctctgctgctggtggtggtggtggttcttctacagtgggggtctctcc 480
Db 421 |||||

```

```
QY 481 ttgagcgtgtacagatggccgctcttctggtacagccgagagagactgtgctggtgtccatg 540
Db 481 ttgagcgtgtacagatggccgctcttctggtacagccgagagagactgtgctggtgtccatg 540
QY 541 aactacccggtgagagccttctgctccctgcccctcccgagagccgagagcccgccg 600
Db 541 aactacccggtgagagccttctgctccctgcccctcccgagagcccgagagcccgccg 600
QY 601 aatgtgggtctctctggtatcagagctggccctgagctggtgagagagccgagccgctcggtg 720
Db 601 aatgtgggtctctctggtatcagagctggccctgagctggtgagagagccgagccgctcggtg 720
QY 721 ggcatacactgctgtcccgcccgccagccgagccgctgttccacagggccgctgtgcagagc 780
Db 721 ggcatacactgctgtcccgcccgccagccgagccgctgttccacagggccgctgtgcagagc 780
QY 781 ggtgcccccaatggaacctgggcccagctgggcatggtggagagcccgctgcagggccacg 840
Db 781 ggtgcccccaatggaacctgggcccagctgggcatggtggagagcccgctgcagggccacg 840
QY 841 cagctggccacactgtggtgctgtctccagggcagctggtggaatgacacagagctg 900
Db 841 cagctggccacactgtggtgctgtctccagggcagctggtggaatgacacagagctg 900
QY 901 gtagcctgcttcggacagacagcagcgagctcctgtgaaacacgaatggcagctgctg 960
Db 901 gtagcctgcttcggacagacagcagcgagctcctgtgaaacacgaatggcagctgctg 960
QY 961 cctcaagaagcgtctctccggttctctcctgctcctggtgtagatggagacttccctcagt 1020
Db 961 cctcaagaagcgtctctccggttctctcctgctcctggtgtagatggagacttccctcagt 1020
QY 1021 gacaccccagagccctatcaacgagggagacttccagggcctgaggtgctggtgggt 1080
Db 1021 gacaccccagagccctatcaacgagggagacttccagggcctgaggtgctggtgggt 1080
QY 1081 gtggtgaagatgagggctcgtatttctggttttagggggcccgagcttcagcaagac 1140
Db 1081 gtggtgaagatgagggctcgtatttctggttttagggggcccgagcttcagcaagac 1140
QY 1141 aacgagctctctatcagccggccgagcttctggtccggggtgcgggttcccccag 1200
Db 1141 aacgagctctctatcagccggccgagcttctggtccggggtgcgggttcccccag 1200
QY 1201 gtaagtgaactggcagcagagcgtgtgctcctgcatatcacagactggctgcatcccgag 1260
Db 1201 gtaagtgaactggcagcagagcgtgtgctcctgcatatcacagactggctgcatcccgag 1260
QY 1261 gaccggcagcctgagggagggccctgagcgtggtggtggcgaccacacaaatgctgtgc 1320
Db 1261 gaccggcagcctgagggagggccctgagcgtggtggtggcgaccacacaaatgctgtgc 1320
QY 1321 cccgtggccagctggtggtggagctggtgcccaggggtcccggttctacgctacgtc 1380
Db 1321 cccgtggccagctggtggtggagctggtgcccaggggtcccggttctacgctacgtc 1380
QY 1381 ttggaacacgctgtctccagctctcctgcccctgtgagtggtggtgcccacggctac 1440
Db 1381 ttggaacacgctgtctccagctctcctgcccctgtgagtggtggtgcccacggctac 1440
QY 1441 gagatgagttcatctttggatcccccttgaccctctcgaactacacagcgagagag 1500
Db 1441 gagatgagttcatctttggatcccccttgaccctctcgaactacacagcgagagag 1500
QY 1501 aaaaatttcgcccagcagctgactgactggtggccaacttttcccgacagggggtatccc 1560
Db 1501 aaaaatttcgcccagcagctgactgactggtggccaacttttcccgacagggggtatccc 1560
QY 1561 aatgagccccgagagcccccaagggcccaaatggcccccggtacacggcggtggtcagcag 1620
```

```
Db 1561 aatgagccccgagagcccccaagggcccaaatggcccccggtacacggcggtggtcagcag 1620
QY 1621 tagcttagtctgacctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1680
Db 1621 tagcttagtctgacctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1680
QY 1681 gccttctggaacgcttctctctctctctctctctctctctctctctctctctctctct 1722
Db 1681 gccttctggaacgcttctctctctctctctctctctctctctctctctctctctctct 1722

RESULT 3
AAQ99002
ID AAQ99002 standard; DNA; 2256 BP.
XX
AC AAQ99002;
XX
DT 31-MAR-1996 (first entry)
XX
DE Human acetylcholinesterase (AChE) gene.
XX
KW Acetylcholinesterase; acetyl cholinesterase; EC-3.1.1.7;
KW chromosome-7q22; acetylcholine-hydrolyzing enzyme; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 160..2206
FT /*tag= a
XX
PN W09523158-A1.
XX
PD 31-AUG-1995.
XX
PF 28-FEB-1995; 95WO-US02806.
XX
PR 09-JAN-1995; 95US-0370156.
PR 28-FEB-1994; 94US-0202755.
XX
PA (KOHN/) KOHN K I.
PA (YISS ) YISSUM RES & DEV CO.
PI Shanl M, Soreq H, Zakut H;
XX
DR WPI; 1995-311499/40.
DR P-PSDB; AAR80726.
XX
PT Alternative forms of human acetyl cholinesterase (ChE) gene -
PT expressed in transgenic animal assay system for evaluating anti-ChE
PT activity of organo:phosphate(s), etc. or as model of ChE imbalance
XX
PS ClalM 3; Fig.1A; 55pp; English.
XX
CC This DNA sequence encoding human acetylcholinesterase is useful
CC for producing transgenic animals which express AChE. The
CC transgenic animals are in turn useful as an assay system for
CC determining the anti-ChE activity of organophosphates, carbamates,
CC anti-ChE drugs, plant glycoalkaloids and snake venoms. This gene
CC contains a promoter region, 6 exons (E1 to E6) and 4 introns (I1-
CC I4).
XX
SQ Sequence 2256 BP; 390 A; 757 C; 680 G; 429 T; 0 other;
```

```
Query Match 99.8%; Score 1722; DB 16; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagcccccgagctgtgctgacacgcttccctggttccccactcttctctc 60
Db 160 atgagcccccgagctgtgctgacacgcttccctggttccccactcttctctc 219
```


QY 61 ctccctgactcctggtgagagagtgagggtcgagggccggaggatcgagagctgctg 120
|||||
Db 220 ctccctgctcctggtgagagagtgagggtcgagggccggaggatcgagagctgctg 279
|||||
QY 121 gtacggtgcgtggggccggtcgcggggcattcgctgaagaccccccggggccctgtc 180
|||||
Db 280 gtacggtgcgtggggccggtcgcggggcattcgctgaagaccccccggggccctgtc 339
|||||
QY 181 tctgcttctcctgggacatccccctttgcggagccaccctatgggaccccgctcttctgcca 240
|||||
Db 340 tctgcttctcctgggacatccccctttgcggagccaccctatgggaccccgctcttctgcca 399
|||||
QY 241 ccggagcccaagacgcttggtcaggggtggtagacgctacaacattccagagtctgc 300
|||||
Db 400 ccggagcccaagacgcttggtcaggggtggttagacgctacaacattccagagtctgc 459
|||||
QY 301 taccatatgtgacacccctataccacaggttttgagggcaccagagatgtggaaccccaac 360
|||||
Db 460 taccatatgtgacacccctataccacaggttttgagggcaccagagatgtggaaccccaac 519
|||||
QY 361 cgtgagctgagcagagactgctgtacctcaacgctgtggacacatacccccgcttaca 420
|||||
Db 520 cgtgagctgagcagagactgctgtacctcaacgctgtggacacatacccccgcttaca 579
|||||
QY 421 tccccacccctcctcgtctgagatctatgggggtaggctcttacagtggggctcctcc 480
|||||
Db 580 tccccacccctcctcgtctgagatctatgggggtaggctcttacagtggggctcctcc 639
|||||
QY 481 ttggagcgtgtacgatggcgctctctgtgtacagggccgagagactgctggtgtccatg 540
|||||
Db 640 ttggagcgtgtacgatggcgctctctgtgtacagggccgagagactgctggtgtccatg 699
|||||
QY 541 aactacccgggtggagcctttgctcctcctgcccctgcccggggccgagagggcccgggc 600
|||||
Db 700 aactacccgggtggagcctttgctcctcctgcccctgcccggggccgagagggcccgggc 759
|||||
QY 601 aatgtgggtctcctgatacagagctggccctcagtggtgcagagacgtggcagcc 660
|||||
Db 760 aatgtgggtctcctgatacagagctggccctcagtggtgcagagacgtggcagcc 819
|||||
QY 661 ttccgggggtgacccgacatcagtgagcgtgtttggggagagcgcgggagccgctcggtg 720
|||||
Db 820 ttccgggggtgacccgacatcagtgagcgtgtttggggagagcgcgggagccgctcggtg 879
|||||
QY 721 ggcatacactgtgtcccccgcagccggggcctgttccacagggccgtgctgcagagc 780
|||||
Db 880 ggcatacactgtgtcccccgcagccggggcctgttccacagggccgtgctgcagagc 939
|||||
QY 781 ggtgcccccaatggacccctggccacggtgggcatgggagagcccgctcgagggccacg 840
|||||
Db 940 ggtgcccccaatggacccctggccacggtgggcatgggagagcccgctcgagggccacg 999
|||||
QY 841 cagctggcccaactgtgggctgtctccagggcgccactggtgggaatgacacagagctg 900
|||||
Db 1000 cagctggcccaactgtgggctgtctccagggcgccactggtgggaatgacacagagctg 1059
|||||
QY 901 gtacgctccttgagacacgacagcgaggtcctggtgaaccacgaatggacgctgctg 960
|||||
Db 1060 gtacgctccttgagacacgacagcgaggtcctggtgaaccacgaatggacgctgctg 1119
|||||
QY 961 cctcaagaaagcgtcttcgggttctcctcgttcctggtggttagatggagacttccctcagt 1020
|||||
Db 1120 cctcaagaaagcgtcttcgggttctcctcgttcctggtggttagatggagacttccctcagt 1179
|||||
QY 1021 gacacccagagccctcatcaacgcggggagagacttccacggcctgcaggtgctggtgggt 1080
|||||
Db 1180 gacacccagagccctcatcaacgcgggagagacttccacggcctgcaggtgctggtgggt 1239
|||||
QY 1081 gtggtgaagagatgagggctcgtatcttctggtttacggggggcccccaggttccagaaagc 1140
|||||
Db 1240 gtggtgaagagatgagggctcgtatcttctggtttacggggggcccccaggttccagaaagc 1299
|||||
QY 1141 aacgagtctctcatcagccgggcgagttccttgccgggggtgcgggtcggggttccccag 1200
|||||

Db 1300 aacgagtctctcatcagccgggcgagttccttgccgggggtgcgggttccccag 1359
|||||
QY 1201 gtaagtgaacctggcagccgaggtgtggtcctgcattacacagactggtgcaccccgag 1260
|||||
Db 1360 gtaagtgaacctggcagccgaggtgtggtcctgcattacacagactggtgcaccccgag 1419
|||||
QY 1261 gacccggacgcttgagggggagccctgagcagatggtggtgggacacacaatgtcgtgctc 1320
|||||
Db 1420 gacccggacgcttgaggggagccctgagcagatggtggtgggacacacaatgtcgtgctc 1479
|||||
QY 1321 ccgctggccagagctggtgggcagctggtcccaggtgctcccaggttcccgggttctacgctacgctc 1380
|||||
Db 1480 ccgctggccagagctggtgggcagctggtcccaggttcccgggttctacgctacgctc 1539
|||||
QY 1381 ttggaacacgctgcttccacgctcctctgcccctgtggatgggggtgccccacggctac 1440
|||||
Db 1540 ttggaacacgctgcttccacgctcctctgcccctgtggatgggggtgccccacggctac 1599
|||||
QY 1441 gagatcaggttcattctttgggatacccccttgagacccctctcgaaactacacgggcagagag 1500
|||||
Db 1600 gagatcaggttcattctttgggatacccccttgagacccctctcgaaactacacgggcagagag 1659
|||||
QY 1501 aaaaactctcccgagcagctgagtcgatactgggccaactttgcccgcacaggggatccc 1560
|||||
Db 1660 aaaaactctcccgagcagctgagtcgatactgggccaactttgcccgcacaggggatccc 1719
|||||
QY 1561 aatgagcccgagaccccaaggcccccaaatggccccgtacacggcggggtccagcag 1620
|||||
Db 1720 aatgagcccgagaccccaaggcccccaaatggccccgtacacggcggggtccagcag 1779
|||||
QY 1621 tagcttagctgacactgcccgcgctgagagtgcgcgggggctgcgcgccagggcctgc 1680
|||||
Db 1780 tagcttagctgacactgcccgcgctgagagtgcgcgggggctgcgcgccagggcctgc 1839
|||||
QY 1681 gcctctggaacgcttccctcccacaaattgctcagcgcaccc 1722
|||||
Db 1840 gcctctggaacgcttccctcccacaaattgctcagcgcaccc 1881
|||||
RESULT 4
AAS17547
ID AAS17547 standard; DNA; 5767 BP.
XX
AC AAS17547;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM034.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
transgenic plant; acetylcholinesterase poisoning; chemical warfare;
muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (NASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, SoREQ H, Arntzen C, Mason H;

|||||
Db 2398 aatgagcccgagagcccaagcccaacaaatggccccctacacggcgggggctcagcag 2457
QY 1621 tacgttagctgagactgcgcgcctggaggtgcgcgggggctgcgcgccagcgcctgc 1680
Db 2458 tacgttagctgagactgcgcgcctggaggtgcgcgggggctgcgcgccagcgcctgc 2517
QY 1681 gctcttggaacgccttctcccaaatgtctcagcgcgcacactga 1725
Db 2518 gctcttggaacgccttctcccaaatgtctcagcgcgcacactga 2562
RESULT 5
AAS17548
ID AAS17548 standard; DNA; 14446 BP.
XX AC AAS17548;
XX DT 25-FEB-2002 (first entry)
XX DE Plasmid pTM036.
XX Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM036;
KW circular; ds.
XX OS Synthetic.
XX PN WO200171014-A2.
XX PD 27-SEP-2001.
XX PF 16-MAR-2001; 2001WO-US08468.
XX PR 17-MAR-2000; 2000US-190440P.
XX PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASON/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX Claim 11; Page 32-41; 42pp; English.
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable

CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM036, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;
SQ

Query Match 99.6%; Score 1718.6; DB 24; Length 14446;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1721; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 atgaggccccgcagtgctgtgcacagccttccctggcttcccaactcttctctc 60
Db 142 atggtccccgcagtgctgtgcacagccttccctggcttcccaactcttctc 201
QY 61 ctccctcgtcctcgtgagagagtgagggtcggggcgagggcgagggatgcagagctg 120
Db 202 ctccctcgtcctcgtgagagagtgagggtcggggcgagggcgagggatgcagagctg 261
QY 121 gtgacggtcgtggggcgcggtgcggggtccttcgctcctcctcctcctcctcctc 180
Db 262 gtgacggtcgtggggcgcggtgcggggtccttcgctcctcctcctcctcctcctc 321
QY 181 tctgcttctcctgggcatccctcttgcggagccaccctcctggaccctcgtcttctgcca 240
Db 322 tctgcttctcctggcctccctcttgcggagccaccctcctggaccctcgtcttctgcca 381
QY 241 cggagcccaagcagccttggtcaggggtgtagagcgtacacaccttccagatgtctgc 300
Db 382 cggagcccaagcagccttggtcaggggtgtagagcgtacacaccttccagatgtctgc 441
QY 301 taccaatatgtgacacacctataccaggttttgaggggcacccagatgtgagaccccaac 360
Db 442 taccaatatgtgacacacctataccaggttttgaggggcacccagatgtgagaccccaac 501
QY 361 cgtgagctgagcgagagactgcctgtacctcaacgctgtgacacacatacccccgccctaca 420
Db 502 cgtgagctgagcgagagactgcctgtacctcaacgctgtgacacacatacccccgccctaca 561
QY 421 tcccccaacctcctcctcgtcgtatctatgggggtggtcttctacagtggggctcctcc 480
Db 562 tcccccaacctcctcctcgtcgtatctatgggggtggtcttctacagtggggctcctcc 621
QY 481 ttggacgtgtacgagtcggcgtcttctgtacagggcgagagagactgtgctggtctccatg 540
Db 622 ttggacgtgtacgagtcggcgtcttctgtacagggcgagagagactgtgctggtctccatg 681
QY 541 aactaccgggtggagccttcttgcctcctcctcctcctcctcctcctcctcctcctc 600
Db 682 aactaccgggtggagccttcttgcctcctcctcctcctcctcctcctcctcctcctc 741
QY 601 aatgtgggtcctcgtatcagagagcgtggcctcagtgagtgaggagaacgtgagcagcc 660
Db 742 aatgtgggtcctcgtatcagagagcgtggcctcagtgagtgaggagaacgtgagcagcc 801
QY 661 ttccggggtgacccgacatcagtgacgctgtttgggagagcgcgaggagcgcctcgggtg 720
Db 802 ttccggggtgacccgacatcagtgacgctgtttgggagagcgcgaggagcgcctcgggtg 861
QY 721 ggcattcacctcgtgtcccccgcagcccggggcttctccacaggggcgtgtctgcagagc 780
Db 862 ggcattcacctcgtgtcccccgcagcccggggcttctccacaggggcgtgtctgcagagc 921
QY 781 ggtgcctcccaatgacaccttggccacaggtggcatgggagagggccctgcagggccacg 840
Db 922 ggtgcctcccaatgacaccttggccacaggtggcatgggagagggccctgcagggccacg 981
QY 841 cagctggcccaaccttggggtctcctccagcgcgacgtggtgggaatgacacagagctg 900
Db 982 cagctggcccaaccttggggtctcctccagcgcgacgtggtgggaatgacacagagctg 1041

Qy	901	gtagctgctcttgagacagacagcgaggtctctctggtgaaccacgaaatggcacgtgctg	960
Db	1042	gtagctgctctcgagacagacagcgaggtctctctggtgaaccacgaaatggcacgtgctg	1101
Qy	961	cctcaagaagcgctctccggtctctctctgctgctggttagatgagacttccctcagt	1020
Db	1102	cctcaagaagcgctctccggtctctctctgctgctggttagatgagacttccctcagt	1161
Qy	1021	gacacccagagagcctcatcaacgcgagagacttccacgcttcagaggtgcgtgggtgggt	1080
Db	1162	gacacccagagagcctcatcaacgcgagagacttccacgcttcagaggtgcgtgggtgggt	1221
Qy	1081	gtggtgaagatgagggctcgattttctgttttacggggcccccagagcttcagcaaaagac	1140
Db	1222	gtggtgaagatgagggctcgattttctgttttacggggcccccagagcttcagcaaaagac	1281
Qy	1141	aacgagctctcatcagccgggcccaggttctctggcccggggtgcggggtcccccag	1200
Db	1282	aacgagctctcatcagccgggcccaggttctctggcccggggtgcggggtcccccag	1341
Qy	1201	gtaagtgaacctggcagccgagggctgtgtctctgcattacagagactggctgcatacccgag	1260
Db	1342	gtaagtgaacctggcagccgagggctgtgtctctgcattacagagactggctgcatacccgag	1401
Qy	1261	gaacccggcacgcttaagaggagccctagacatgtgtgtgggacacaaatgctggtgc	1320
Db	1402	gaacccggcacgcttaagaggagccctagacatgtgtgtgggacacaaatgctggtgc	1461
Qy	1321	cccgtggccagctggctggggagctggctgccaggggtcccgggtctacgctacgtc	1380
Db	1462	cccgtggccagctggctggggagctggctgccaggggtcccgggtctacgctacgtc	1521
Qy	1381	tttgaaacacggtgcttcaacgctctctctggccctgttgatggggggtccccacggctac	1440
Db	1522	tttgaaacacggtgcttcaacgctctctctggccctgttgatggggggtccccacggctac	1581
Qy	1441	gagatcgagttcatctttgggataccccctgacccccctctcgaaactacacggcagagagg	1500
Db	1582	gagatcgagttcatctttgggataccccctgacccccctctcgaaactacacggcagagagg	1641
Qy	1501	aaaaatttcgcccagcagactgatgcgatacttggggccaaactttgccgcacaggggataccc	1560
Db	1642	aaaaatttcgcccagcagactgatgcgatacttggggccaaactttgccgcacaggggataccc	1701
Qy	1561	aatgaagcccgagaccccaaggcccaaaatggcccccggtacacggcggggggtcagcag	1620
Db	1702	aatgaagcccgagaccccaaggcccaaaatggcccccggtacacggcggggggtcagcag	1761
Qy	1621	tacgttagtcttgaaactcgccgtcgaggtgcggcgggggtgcgcgcgccagggcctgc	1680
Db	1762	tacgttagtcttgaaactcgccgtcgaggtgcggcgggggtgcgcgcgccagggcctgc	1821
Qy	1681	gcctctgaaacgcttctccccaaaattgctcagcgccaaactga	1725
Db	1822	gcctctgaaacgcttctccccaaaattgctcagcgactactga	1866

RESULT 6

RESOLU
AAO05998

AAQ05998
ID AAQ05998 standard; DNA: 2253 BP.

XX
DT
060507ZAA
AC
AAQ05998;XX
XX
XXXXXX

DT 16-JAN-1991 (flrst entry)

XX

DE Sequence encoding human acetylcholinesterase (hAChE).

XX

KW Organophosphorous poisoning; OP; cancer; leukaemia;

KW megakaryocytopoiesis; ovarian cancer; ds.

XX

OS Homo sapiens.

XX
c c c c c c c

PN EP388906-A.

QY 601 aatgtgggtctcttgatcagaggtgcccctgagtgaggtgcaagagaacgtgagacc 660
Db 757 aatgtgggtctcttgatcagaggtgcccctgagtgaggtgcaagagaacgtgagacc 816
QY 661 ttcgggggtgaccgacatcagtagcagctgtttggggagagcgcgagcgccctcggtg 720
Db 817 ttcgggggtgaccgacatcagtagcagctgtttggggagagcgcgagcgccctcggtg 876
QY 721 ggcatagcacctgctgtcccccgcagccgagcgccgtgttccacagggccgtgctgcagagc 780
Db 877 ggcatagcacctgctgtcccccgcagccgagcgccgtgttccacagggccgtgctgcagagc 936
QY 781 ggtgcccccaatgacacctgagccacgtagggcagtagggagagggccgtgcagagggccacg 840
Db 937 ggtgcccccaatgacacctgagccacgtagggcagtagggcagtagggccgtgcagagggccacg 996
QY 841 cagctggccccacctgtgtggctgtctccacagcgagcagctgtgggaatgacacagagctg 900
Db 997 cagctggccccacctgtgtggctgtctccacagcgagcagctgtgggaatgacacagagctg 1056
QY 901 gtacgctccttcggagacagccagcaggtcctgtgtgaaccacgaatggcacgtgctg 960
Db 1057 gtacgctccttcggagacagccagcaggtcctgtgtgaaccacgaatggcacgtgctg 1116
QY 961 cctcaagaaagcgtctccgtctctcctgtctcctgtgctgtgtagatgagaccttctcagt 1020
Db 1117 cctcaagaaagcgtctcctgtctcctgtgctgtgtagatgagaccttctcagt 1176
QY 1021 gacacccagagggcctcatcaacgagggagagacttccacggcctcaggtgctgtggt 1080
Db 1177 gacacccagagggcctcatcaacgagggagagacttccacggcctcaggtgctgtggt 1236
QY 1081 gtggtgaaggatgagggctgtattttctgtttacggggccccaggttcagcaaaagac 1140
Db 1237 gtggtgaaggatgagggctgtattttctgtttacggggccccaggttcagcaaaagac 1296
QY 1141 aacagatctctcatcagcggcgagctcctgctgaggggtgaggggtgaggggttccccag 1200
Db 1297 aacagatctctcatcagcggcgagctcctgctgaggggtgaggggtgaggggttccccag 1356
QY 1201 gtaagtacctggagcggagctgtgtctgctgattacacagactgctgctacatccag 1260
Db 1357 gtaagtacctggagcggagctgtgtctgctgattacacagactgctgctacatccag 1416
QY 1261 gaccggcacgctgagggagggccctgagcagatgtggtggtggcgacccaatgtctgtgc 1320
Db 1417 gaccggcacgctgagggagggccctgagcagatgtggtggtggcgacccaatgtctgtgc 1476
QY 1321 cccgtggccccagctggtggcgagctgctgcccaggtgcccggtgtacgctacatc 1380
Db 1477 cccgtggccccagctggtggcgagctgctgcccaggtgcccggtgtacgctacatc 1536
QY 1381 ttggaacacctgtctccacgctctcctggtccctgtggtggtgggtgcccacggctac 1440
Db 1537 ttggaacacctgtctccacgctctcctggtccctgtggtggtgcccacggctac 1596
QY 1441 gagatcgagttcatcttgggatccccctgagcccccttcgaaactacacggcgagagag 1500
Db 1597 gagatcgagttcatcttgggatccccctgagcccccttcgaaactacacggcgagagag 1656
QY 1501 aaaaacttcgccccagcagctgactgcgatactggggcgaactttgcccgcacagggatccc 1560
Db 1657 aaaaacttcgccccagcagctgactgcgatactggggcgaactttgcccgcacagggatccc 1716
QY 1561 aatgagccccgagacccccagggcccaaatggtgccccgtacacggcggggtcagcag 1620
Db 1717 aatgagccccgagacccccagggcccaaatggtgccccgtacacggcggggtcagcag 1776
QY 1621 tacgttagctggaacctgagccgctgaggtgaggtgcggcgggggtcgcgcgccagccctgc 1680
Db 1777 tacgttagctggaacctgagccgctgaggtgaggtgaggtgcggcgggggtcgcgcgccagccctgc 1836
QY 1681 gcttcttggaacccgttctctcccccaaatgtctcagcgccacc 1722

Db 1837 gcttcttggaacccgttctctcccccaaatgtctcagcgccacc 1878

RESULT 7
AAQ05999

ID AAQ05999 standard; DNA; 1800 BP.

XX AAQ05999;

XX 16-JAN-1991 (first entry)

XX Sequence encoding foetal human acetylcholinesterase (hAChE).

XX Organophosphorous poisoning; OP; cancer; leukaemia;

XX megakaryocytopoiesis; ovarian cancer; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1500

XX 3'UTR /*tag= a

XX 1501..1800

XX /*tag= b

XX EP388906-A.

XX 26-SEP-1990.

XX 20-MAR-1990; 90EP-0105274.

XX 21-MAR-1989; 89IL-0089703.

XX (YISS) YISSUM RES DEV CO.

XX Soreq H, Zakut H;

XX WPI; 1990-291865/39.

XX P-PSDB; AAR06990.

XX Human acetylcholinesterase DNA and prodn. of recombinant hAChE -

XX for treatment of organo-phosphorous poisoning and diagnosis of haemo-

XX cytopoeitic disorders and ovarian carcinomas.

XX Disclosure; Fig 1c; 47pp; English.

XX hAChE is useful as an active pharmacological component for the

XX prophylaxis and treatment of organophosphorous poisoning, and

XX post-surgical apnea due to succinylcholine administration.

XX CDNA probe to the sequence may be used in diagnosis of various

XX leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.

XX Bases given as N are obscured in the patent specification.

XX Sequence 1800 BP; 330 A; 602 C; 539 G; 326 T; 3 other;

Query Match 76.2%; Score 1313.6; DB 11; Length 1800;

Best Local Similarity 98.4%; Pred. NO. 5.6e-268;

Matches 1358; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

QY 346 atgtggaaccccaacacgctgagctgagctgagcgagagactcctgtacctcaacgtgtggacacca 405

Db 1 atgtggaaccccaacacgctgagctgagcgagagactcctgtacctcaacgtgtggacacca 60

QY 406 tacccccggcctacatccccacccctgtcctcgtctggtatgtatgggtggtctctac 465

Db 61 tacccccggcctacatccccacccctgtcctcgtctggtatgtatgggtggtctctac 120

QY 466 agtggggcctcctcttgagctgtacgatggccgctctcttggtagagggagagact 525

Db 121 agtggggcctcctccttgagctgtacgatggccgctctcttggtagagggagagact 180

QY 526 gtgctggtgtccatgaactacacgggtgtggagcctttgcttctgcccgtgcggggagc 585

Db 181 gtcttggtgccccgaactaccgggtgggagcctttggctctcttgccctgcgcgggagc 240
QY 586 cgagagccccgggcaatgtgggtctcttgatcacagaggtggccctgcagtggtgacag 645
Db 241 cgagagccccgggcaatgtgggtctcttgatcacagaggtggccctgcagtggtgacag 300
QY 646 gagacgtggcagccttcgggtgggtgaccogacatcagtgacgtgttttggggagagcg 705
Db 301 gagacgtggcagccttcgggtgggtgaccogacatcagtgacgtgttttggggagagccg 360
QY 706 gggagcgcctcgtgggtgacgtgacacgtctgtcccgcccaagcgggcccgtttccacagg 765
Db 361 gggagcgcctcgtgggtgacgtgacacgtctgtcccgcccaagcgggcccgtttccacagg 420
QY 766 gccgtgctgcagagcggtgcccccaatggaccctgggcccacggtgggcatgggagagccc 825
Db 421 gccgtgctgcagagcggtgcccccaatggaccctgggcccacggtgggcatgggagagccc 480
QY 826 cgctcgagggccacgcagctggcccacctgtgggtgtctctccagcggcactggtggg 885
Db 481 cgctcgagggccacgcagctggcccacctgtgggtgtctctccagcggcactggtggg 540
QY 886 aatgacacagagctgttagctccttcctggacacgaccagcgaggtcctggtgaaccac 945
Db 541 aatgacacagagctgttagctccttcctggacacgaccagcgaggtcctggtgaaccac 600
QY 946 gaatggcagctgtgctcctcaagaaagcgtcttcctggtctctcctgctggtgtagat 1005
Db 601 gaatggcagctgtgctcctcaagaaagcgtcttcctggtctctcctgctggtgtagat 660
QY 1006 gggagctctcagtgacaccccgagggccctcatcaaacgcgggagacttccacggcctg 1065
Db 661 gggagctctcagtgacaccccgagggccctcatcaaacgcgagacttccacggcctg 720
QY 1066 caggctcgtggtgtggtgaagatgaggtcgtattttctgtttacggggccca 1125
Db 721 caggctcgtggtgtggtgaagatgaggtcgtattttctgtttacggggccca 780
QY 1126 ggttcagcaagaacacagtagtctctcatcagccggcgagttcctggccgggtgctgg 1185
Db 781 ggttcagcaagaacacagtagtctctcatcagccggcgagttcctggccgggtgctgg 840
QY 1186 gtccgggttccccaggttaagtaccctggcagccgaggtgtgtctctgcattacacagac 1245
Db 841 ctccgggttccccaggttaagtaccctggcagccgagctgtgtctctccattacacagac 900
QY 1246 tggctcatcccgagaccccgacacgctgagggagggccctgagcgtgtgtggggcag 1305
Db 901 tggctcatcccgagaccccgacacgctgagggagggccctgagcgtgtgtggagcag 960
QY 1306 cacaatgtcgtgtgcccggtggcccagctggtggggc-gactggctgccaggggtgccc 1364
Db 961 cacaatgtcgtgtgcccggtggcccagctggtggggcagactggtgccaggggtgcccg 1020
QY 1365 ggt-ctacgctc-actgtttgaacaccgtgtcttccacgctctcctggccctgtgtagt 1422
Db 1021 ggtactacgctcaactgtttgaacaccgtgtcttccacgctctcctggccctgtgtagt 1080
QY 1423 ggggtgccccaggtcacagatcgagttcattttggggtacccctggacccctctcga 1482
Db 1081 cgggtgtcccccaggtcacagatcgagttcattttggggtacccctggacccctctcga 1140
QY 1483 aactacagcgagagagaaaaattctcggcccagcgactgatgcgactggtggccaacttt 1542
Db 1141 aactacagcgagagagaaaaattctcggcccagcgactgatgcgactggtggccaacttt 1200
QY 1543 gcccgacaggggatcccaatgagccccagagaccccaagcccccaaatggcccccgtag 1602
Db 1201 gcccgacaggggatcccaatgagccccagagaccccaagcccccaaatggcccccgtag 1260
QY 1603 acggggggggtcagcagtagcttagtctgagcctcgccgctgaggtgctggcggggg 1662

Db 1261 acccgccccgggctcagcagtagcttagtctggaacctggcgccgtggaggtgcgcggggg 1320
QY 1663 ctgcgcgccagcctgcgccttcttggaaccttctcccccataattgtcagcgccacc 1722
Db 1321 ctgcgcgccagcctgcgccttcttggaaccttctcccccataattgtcagcgccacc 1380
RESULT 8
ABI99790
ID ABI99790 standard; cDNA; 2089 BP.
XX
AC ABI99790;
XX
DT 07-MAR-2002 (first entry)
XX
Mouse ischaemic condition related cDNA sequence SEQ ID NO:890.
XX
DE Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO20018188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
WI: 2002-034733/04.
XX
Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 2244-2245; 2690pp; English.
XX
The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC the expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression levels of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 2089 BP; 409 A; 638 C; 596 G; 446 T; 0 other;

Query Match 74.3%; Score 1282; DB 24; Length 2089;
Best Local Similarity 84.0%; Pred. No. 2.6e-261;
Matches 1447; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1 atagagcccccgagtgctgtgcacacgccttccctggcttcccaactctctctc 60
Db 1 atgagcctcctgggtatccccctgcatacaccttcccttcttccactctctc 60
QY 61 ctctctcgtcctgggtgagagtgaggcgtgagggcgagggagatgcagagctgctg 120
Db 61 ctctctcctcctggagaggggcaaggcgtgagggcgagagaccgagagctgctg 120

Oy 121 gtgacggtgctggggcgccgctgcggggcattccgctcgaagaccccccgggggccctgtc 180
Db 121 gtgaggggtcagagggggcagctgaggggcatcccgccgaagggccctggagggccagtc 180
Oy 181 tctgcttctccctgggcatccctcttgcggagccaccacatgggagcccgctgcctttctgcca 240
Db 181 tcaagctttctgggcatccctcttgcagagccaccctggggctcagctagatttatgcca 240
Oy 241 ccgagagcccaagcagccttggttcagggggtggtagacgtatacaaccttccagagtctgtc 300
Db 241 ccagagcccaagcggccctgggtcagagagtttggtgctaccaccttccaaaatgtctgc 300
Oy 301 taccaatatgtgacacccctataccaggttttgaggccaccgagatgtggaaccccaac 360
Db 301 taccagtcagtggaaccccttaccctgggttgagggtactgagatgtggaaccccaac 360
Oy 361 cgtgagctgagcagagactgcctgtacatcaacgtgtggaacccatacccccgcctaca 420
Db 361 cgagagttgagtggaactgcctgtatcttaatgtgtggaacccatacccccagactgct 420
Oy 421 tccccacccctgtcctcgtctggtatcatatgggggtggtcttatacagtggggcctctcc 480
Db 421 tctccacacctgtcctcatctggtatcatatgggggtggtttctacagcggagcgccctcc 480
Oy 481 ttgacgtgtacgatggcgcttctgtgtacagccgagagagactgtgctgtgtccatg 540
Db 481 ttggatgtgtatgagcgccgtttctgtggccaggtgtgaggagctgtgtgttatctatg 540
Oy 541 aactacccgggtgggagcctttggcttctgcctgcggggagcggagagggcccccgggc 600
Db 541 aactacccagtggaacctttggcttctgttggcctaccaggaagcagagaaagccctggc 600
Oy 601 aatgtgggtctcctggatcaagagggtggccctgcagtgagggtggcgagaaacgttggcagcc 660
Db 601 aatgtagtgctgtgatacaacgggtgccttgcattgcaatgggtgcaagaaatattgcagcc 660
Oy 661 ttgggggtgaccagacactagtcagctgtgtttggggagagcgggagccgctcgatg 720
Db 661 ttggggggcccgagatgtcagtgactctgtttggggagagtgcggtgcagccctccgtg 720
Oy 721 ggcatagcacctgtctcccgcccgagcgggctgttccacagggccgctgctgcagagc 780
Db 721 ggcatagcataactgtccctgcgcagcaggagccttctccacagggctgtctccagagt 780
Oy 781 ggtgcccccaatggacctgggcccagctgggcatgggagagggcccgctgcagggccacg 840
Db 781 ggcacacccaatggccctgggcccactgtgagtgctgagagggccagggcgagggccaca 840
Oy 841 cagctggcccacctgtgtggctgtcctccagcggcagctgtgtggaatgacacagagctg 900
Db 841 ctgtcggcccgcctgtgtgggtgtccccccagggtggcgtgtgtggcaatgacaccgagctg 900
Oy 901 gtgacctgcttcggagacagcagcagctgctgttggtgaaccagaaatggcacgtgctg 960
Db 901 atagcctgtctgaggaacaggcccgctcaggacctgtgtggaaccagagtggcacgtctcg 960
Oy 961 cctcaagaaagcgttctccggttctcctcgtcctcgtgtggttagatggagacttctcagtt 1020
Db 961 cctcaagaaagtattctccgatttctcctcgtgtcctgtgtgtagacggggagacttctcagtt 1020
Oy 1021 gacacccagagggccctcatcaacgcggagagacttccacggcctcagctgctgtgtggt 1080
Db 1021 gacacccggaggccctcatcaataactggagatttccaagacctcaggtgtgtgtggt 1080
Oy 1081 gtggtagaaggatgagggctcgtattttctgtttacggggggcccgaggttccagcaagac 1140
Db 1081 gtggtagaaggacgagggctcctacttctgtgtttacgggggttcccgaggttccagaaagac 1140
Oy 1141 aacgagttctctatcagcggggccaggttctcgtggccgggtgtcgggttcgggttcccccag 1200
Db 1141 aatgaatctctatcagcggggccaggttctcgtgggtgtcgggtgtcgggttaaccccaa 1200
Oy 1201 gtaagtgaactggcagcggaggtgtgtcctcgtcattacacagactggtcgtatcccgag 1260

Db 1201 gcaagtgaacctggcgccgagagctgtgctcctgcattacacagactggtgcacctgag 1260
Oy 1261 gaccgcgacgcctgagggagccctgagcagatgtgggtggcgacacaaatgtcgtgtgc 1320
Db 1261 gacctactcaactagagatgccaatgagtcagtggttaggcgacacaaacgttgtgtgc 1320
Oy 1321 ccctggtgcccagctggtcgggagctggtgcctgcccaggggtgcctggttctacgctcagtc 1380
Db 1321 cctgtggcccagctggtcgggagctggtgcctgcccagggggccgggtctatgctctacatc 1380
Oy 1381 ttgtgaacccgtgtccacgctcctcctggtccctgtggtatgggtgtggtgtgcccacgctac 1440
Db 1381 ttgtgaacccgtgtcctccacactgactgtgcccctcctggtggtgtggtgtgcccctatgctat 1440
Oy 1441 gagatcaggttcatctttggggtacccctggacccctcgcgaactacacacgagagagag 1500
Db 1441 gaaatcaggttcatctttggggtcctcctggtacccctcgtcgtgaactacacacgagagag 1500
Oy 1501 aaaaattctgcccagcagctgagtcgatactggtgccaactttgcgcgacagagggatccc 1560
Db 1501 aggatctttgctcagcagacttatgaaatctggaactgacaaattttgcccgcacaggggacccc 1560
Oy 1561 aatgagccccagacccaagccccacaatgcccccgatcacacggtggggggtcagcag 1620
Db 1561 aatgacctcagagactccaaaatctccacagtggtgccacactgcccgcagcagaa 1620
Oy 1621 tacgttagctgagcctcggcgctggaggtgctggcggggggtcgcgcgcccagggcctgc 1680
Db 1621 tatgtgacctgaacctgaagccttagaggtgcgggggagctgcgcgcccagacactgc 1680
Oy 1681 gctctgtgaacgcttctctcccaaaattgctcagcgccacc 1722
Db 1681 gctctgtgaatcgctttctcccaaaattgctcagcgccacc 1722

RESULT 9
AAS17492 standard; DNA; 12113 BP.
XX
AC AAS17492;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human acetylcholinesterase, AChE, gene.
XX
KW Human; ds; AChE; acetylcholinesterase; polymorphic variant; haplotyping;
KW genotyping; neurological disease; Parkinson's disease;
KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT variation replace (1950,C)
FT /*tag= a
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (2237,T)
FT /*tag= b
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (4044,T)
FT /*tag= c
FT /*standard_name= "Single nucleotide polymorphism"
FT CDS 4095..8113
FT /*tag= d
FT /*product= "ACHE"
FT exon 4095..5162
FT /*tag= e
FT /*number= 2
FT intron 5163..5508
FT /*tag= f
FT /*number= 2
FT variation replace (4130,T)
FT /*tag= g


```

FT variation /standard_name= "Single nucleotide polymorphism"
FT /tag= h replace (4195,A)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= h replace (4277,C)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= i replace (4932,G)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= j replace (4967,T)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= k replace (5126,A)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= l replace (5151,A)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= m replace (5871,T)
FT /tag= n /number= 3
FT /tag= o 5994..6988
FT /number= 3
FT /tag= o replace (5871,T)
FT /tag= p /standard_name= "Single nucleotide polymorphism"
FT /tag= q replace (6831,A)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= r 6989..7158
FT /tag= r /number= 4
FT /tag= s 7159..7991
FT /tag= s /number= 4
FT /tag= t 7992..8113
FT /tag= t /number= 5
FT /tag= u replace (8114,A)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= v replace (8200,A)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= w replace (8227,A)
FT /tag= w /standard_name= "Single nucleotide polymorphism"
FT /tag= x replace (8425,G)
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= x
FT WO200179219-A2.
FT PN
FT XX
FT PD 25-OCT-2001.
FT XX
FT PF 11-APR-2001; 2001WO-US11853.
FT XX
FT PR 14-APR-2000; 2000US-197173P.
FT XX
FT PA (GENA-) GENAISSANCE PHARM INC.
FT PA (KAZE/) KAZEMI A.
FT XX
FT PI Bentivegna SC, Chew A, Choi JY, Koshy B;
FT XX
FT DR WPI: 2002-055248/07.
FT DR P-PSDB; AAU11231.
FT XX
FT PT New polymorphic variants comprising acetylcholinesterase (ACHE)
FT PT isogene, useful in expressing ACHE protein for use in screening for
FT PT candidate drugs to treat diseases related to ACHE activity, e.g.
FT PT neurological diseases or cancer -

```

XX Claim 21; Fig 1; 79pp; English.

XX The invention relates to a polynucleotide comprising a polymorphic
 CC variant of an acetylcholinesterase (ACHE) gene or fragment, protein or
 CC complement, the variant comprising an ACHE isogene defined by a haplotype
 CC selected from haplotypes 1-20 listed in the specification. Also included
 CC are methods for haplotyping and genotyping the ACHE gene of an
 CC individual, a method for predicting a haplotype pair for the ACHE gene of
 CC an individual, a method for identifying an association between a trait
 CC and at least one haplotype or haplotype pair of ACHE gene, recombinant
 CC nonhuman organisms transformed or transfected with the polynucleotide
 CC where the organism expresses ACHE protein encoded by the first
 CC nucleotide sequence or encoded by the polymorphic variant sequence,
 CC an isolated antibody specific for and immunoreactive with ACHE,
 CC a method of screening for drugs targeting the polypeptide contacting ACHE
 CC polymorphic variant with a candidate agent and assaying for binding
 CC activity, a computer system for storing and analysing polymorphism data
 CC for ACHE gene and a genome anthology for ACHE gene which comprises ACHE
 CC isogenes defined by haplotypes 1-20 given in the specification.
 CC The Polymorphisms are useful for studying the biological function of
 CC ACHE as well as in identifying drugs targeting this protein for the
 CC treatment of disorder related to its abnormal expression or function.
 CC The polymorphic variants may also be used in screening for compounds
 CC targeting ACHE to treat a specific condition or disease predicted to be
 CC associated with ACHE activity e.g. neurological diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
 CC and tumours. The ACHE gene maps to human chromosome 7q22. The
 CC present sequence is the ACHE gene.

XX Sequence 12113 BP; 2215 A; 3836 C; 3766 G; 2296 T; 0 other;

Query Match 69.4%; Score 1198; DB 24; Length 12113;
 Best Local Similarity 81.8%; Pred. No. 1.9e-243;
 Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

QY 1 atgagccccccagtgctgtgacacgccttccctggtcttcccaactcttctctc 60
 |||
 Db 4095 atgagccccccagtgctgtgacacgccttccctggtcttcccaactctctc 4154
 QY 61 ctctcttgctctctggtgagagtggtggggtgagggccgggagagtgagagtgctg 120
 |||
 Db 4155 ctctcttgctctctggtgagagtggtggggtgagggccgggagagtgagagtgctg 4214
 QY 121 gtacaggtgctgtggggccggtcggtggtggtggtggtggtggtggtggtggtggtg 180
 |||
 Db 4215 gtacaggtgctgtggggccggtcggtggtggtggtggtggtggtggtggtggtggtg 4274
 QY 181 tctgtcttcttgggtccttcccttgcggagccaccatgggaccccgctcttcttgcca 240
 |||
 Db 4275 tctgtcttcttgggtccttcccttgcggagccaccatgggaccccgctcttcttgcca 4334
 QY 241 ccggagcccaagcagcttggtcaggggtgtagagctacaccccttccagagtgctgc 300
 |||
 Db 4335 ccggagcccaagcagcttggtcaggggtgtagagctacaccccttccagagtgctgc 4394
 QY 301 taccatattgtgacacccctataccaggttttgaggccaccagagatgtggaacccaac 360
 |||
 Db 4395 taccatattgtgacacccctataccaggttttgaggccaccagagatgtggaacccaac 4454
 QY 361 cgtgagctgagcagagactgctgtacctcaacgtgtggacaccatcccccggtctaca 420
 |||
 Db 4455 cgtgagctgagcagagactgctgtacctcaacgtgtggacaccatcccccggtctaca 4514
 QY 421 tccccaccccttctctgctgtggtatctatgggggtggtctctacagtgagggtctctcc 480
 |||
 Db 4515 tccccaccccttctctgctgtggtatctatgggggtggtctctacagtgagggtctctcc 4574
 QY 481 ttgacgtgtacagtgccgctctctgtgtacaggtggtacaggtggtggtggtggtggtg 540
 |||
 Db 4575 ttgacgtgtacagtgccgctctctgtgtacaggtggtacaggtggtggtggtggtggtg 4634

Db	5715	gagggagccctgagcgtgtggtggcgaccacaatgtcgtgtgccccgtggccccagct	5774
Qy	1335	ggctggcgagctagctgcccaggggtgcccggtctacgccctacgtcttttgaaacacgtgc	1394
Db	5775	ggctggcgagctgctgccaggggtgcccggtctacgccctacgtcttttgaaacacgtgc	5834
Qy	1395	tccacgctctctctggccctctggtggtgggtgcccacggctacgagatcgagttcat	1454
Db	5835	tccacgctctctctggccctgtggtggtgggtgcccacggctacgagatcgagttcat	5894
Qy	1455	ctttgggataccccctggaaocctctctcgaaactacacgagaggagaaatcttcgccca	1514
Db	5895	ctttgggataccccctggaaocctctctcgaaactacacgagaggagaaatcttcgccca	5954
Qy	1515	gcgactatcgatactgggccaactttgccgcacaggg	1554
Db	5955	gcgactatcgatactgggccaactttgccgcacaggg	5994
RESULT 10			
AAx86613			
ID	AAx86613 standard; cDNA; 1698 BP.		
XX	AAx86613;		
XX	XX		
XX	XX		
DT	15-OCT-1999 (first entry)		
XX	XX		
DE	cDNA encoding an acetylcholinesterase protein.		
XX	XX		
KW	Acetylcholinesterase; AChE; fusion protein; ligand receptor;		
KW	monomer; ligand detection; marker enzyme; ss.		
XX	XX		
OS	Bungarus fasciatus.		
XX	XX		
FH	Key		
FT	1..1698		
FT	CDS		
FT	Location/Qualifiers		
FT	/*tag a		
FT	/transl_except- (pos: 52..54, aa: Leu)		
FT	/transl_except- (pos: 55..57, aa: Cys)		
FT	/transl_except- (pos: 58..60, aa: Ile)		
FT	/transl_except- (pos: 61..63, aa: Pro)		
FT	/transl_except- (pos: 64..66, aa: Ser)		
FT	/transl_except- (pos: 67..69, aa: Cys)		
FT	/transl_except- (pos: 70..72, aa: Val)		
FT	/transl_except- (pos: 73..75, aa: Ala)		
FT	/transl_except- (pos: 76..78, aa: Val)		
FT	/transl_except- (pos: 79..81, aa: Leu)		
FT	/transl_except- (pos: 82..84, aa: Gln)		
FT	/transl_except- (pos: 85..87, aa: Leu)		
FT	/product= "acetylcholinesterase"		
FT	/note= "no termination codon given"		
XX	XX		
PN	FR2773802-A1.		
XX	XX		
PD	23-JUL-1999.		
XX	XX		
PF	22-JAN-1998; 98FR-0000656.		
XX	XX		
PR	22-JAN-1998; 98FR-0000656.		
XX	XX		
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.		
PA	(INSP) INST PASTEUR.		
XX	XX		
PI	Bon C, Choumet V, Cousin X;		
XX	XX		
DR	WPI; 1999-471239/40.		
DR	P-PSDB; AAY30100.		
XX	XX		
FT	A fusion protein comprising an acetyl cholinesterase and ligand		
PT	receptor, useful for detection of ligands		
XX	XX		
PS	Claim 19; Page 98-99; 114pp; French.		


```
OY 686 cgcctgtttggagagacgcgggagccgctcggtggtggcagtcacactgctgtgtcccgccca 745
Db 737 cctctcttggagaagtgcagagcagctcagttagctgcatgttcttctcctggaa 796
OY 746 gccgggggctgttccacagggcgtgctgcagagcgttgcccccaatggaccctgggcca 805
Db 797 gccattcattgtccacagagccattctgcaaaagtgcattcttaagtctctcttggcgg 856
OY 806 cgggtgggcatgggagagcccgctgcagggccacgcagctggccccaccttggggctgtc 865
Db 857 taacatctctttatgaagctaggaacagaacgttgaacttagctaaatt----- 905
OY 856 ctcacagcggcaactgtgtgggaatgacacagagctggtgagctcctcggaacacacag 925
Db 906 -gactggtgtctctagagagatgagactgaaataatacaagtgtctctagaaataaagatc 964
OY 926 cgcaggtcctggtgaaccacgaatggcagctgctgcctcaagaaagcgtcttccggttct 985
Db 965 cccaagaattctctgaatgaagcatttgtgtccctatgggactccttctgtcagtaa 1024
OY 986 ccttcgtgcctgtggtagatggagacttctcagtcagtcacccccagagccctcatcaacg 1045
Db 1025 acttggctccgacctgagtgtgatttctcactgacatgcagacatatattacttgaac 1084
OY 1046 cgggagacttccacggcctgcaggtgctggtggtggtggtggaagatgagggctgtatt 1105
Db 1085 ttggacaatttaaaaaaacccaga ttttgggtgggtggttaataaagatgaaggagacgtt 1144
OY 1106 ttctggtttacggggcccgcttcacgaacacagagctctctcatcagcgggccc 1165
Db 1145 ttttagtctatggtgctcctggcttcagaaagataaacaatagatcataactagaaaag 1204
OY 1166 agtctcctggccgggtgcgggttcgggggttccccaggttaagtgcctggcagccgggctg 1225
Db 1205 aatttcaggaaggtttaaaaaatttttccaggagtgagtgagtttggaaggaatcca 1264
OY 1226 tggctcctgattacacagactggctgcctcccgagagaccggcagcctgaggaggccc 1285
Db 1285 tcttttccattacacagactgggtgagatgatacagagacctgaaactacogtgaggcct 1324
OY 1286 tgagcagatgtgtgggcgaccacaatgtcgtgtgccccgtggccagctggctgggcgac 1345
Db 1325 tgggtgagtggttggggtattataatttcattatgccttgccttgagttgacctaccaagaagt 1384
OY 1346 tggctgccccaggtgcccgggtctacgcctacgtcttcttgaacacccgtgcttccacgctct 1405
Db 1385 tctcagaatggggaataatgccttttctactattttgaacacgcgactcctccaaacttc 1444
OY 1406 cctggccctctggtatgggggtgccccgcgctacgagatcgagttcattcttgggatcc 1465
Db 1445 cgtggccagaatggatgggagtgatgcatggctatggctatgaaattgtcttctgttttac 1504
OY 1466 cctggagccctctcgaaactacacgggcagagggagaaaattcttcccgagcagactgatgc 1525
Db 1505 cctctggaaaagagagataattacacaaaagccgagggaaaattttagtagatccatagtga 1564
OY 1526 gatactgggccaactttccgcgcagacagggatcccaatgagcccccgagac 1575
Db 1565 aacggtgggcaaatgttgcaaaatgggaatcccaatgagactcagaac 1614
```

RESULT 15

AAN60111

ID AAN60111 standard; cDNA; 2445 BP.

XX

AC AAN60111;

XX

DT 27-JUN-1991 (first entry)

XX

DE Sequence encoding a protein having human cholinesterase (ChE)

DE activity.

XX

KW Organophosphorous poisoning; therapy; prophylaxis; diagnosis;

KW pseudochoolinesterase deficiency; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

CDS

FT 1..1908

FT /*tag= a

FT TATA_signal 2374..2379

FT /*tag= b

XX

PN EP206200-A.

XX

PD 30-DEC-1986.

XX

PF 16-JUN-1986; 86EP-0108189.

XX

PR 18-JUN-1985; 85IL-0075553.

XX

PA (YEDA) YEDA RES & DEV CO LTD.

XX

PI Soreq H;

XX

DR WPI; 1986-340581/52.

XX

PT Human cholinesterase-type proteins - produced by host cell contg.

PT recombinant vector which contains DNA fragment coding for the

PT enzyme

XX

PS Claim 8; Page 38-40; 44pp; English.

XX

CC Human cholinesterase-type proteins can be used for the prophylaxis

CC or for the treatment of the effects of organophosphorous cpd.

CC poisoning. Other uses include clinical

CC detection of

CC pseudochoolinesterase deficiencies,

CC elucidation of the active site

CC topography and the AA sequence of

CC AChE for the development of rapid

CC simple clinical methods to detect

CC poisoning or disease-related

CC changes in ChEs, and for detecting

XX mutations in ChE genes.

SQ

Sequence 2445 BP; 814 A; 432 C; 465 G; 734 T; 0 other;

Query Match 18.7%; Score 322.8; DB 7; Length 2445;

Best Local Similarity 54.6%; Pred. No. 6.5e-59;

Matches 780; Conservative 0; Mismatches 627; Indels 21; Gaps 6;

OY 146 ggggcatctgcctgaagaccccggggcctgtctctgtcttctcctgggcatcccttg 205

Db 227 gagggatgaactgacagtgtttgttggaacggttaacagcctttcttggaatccctatg 286

OY 206 cggagccacccatgggaccccgctcgtcttctgccaccggagcccaagcagccttggtcag 265

Db 287 cacagccacctgtgtgtagacttcgattcacaaagccacagctctctgaccaggtggtctg 346

OY 266 ggggtgtagacgtcaaaccttccagagtgtctgtacaaatattgttggaacccctatacc 325

Db 347 atatttgactgcccaaaaatgcaaatctgtgtctgcagacaatagatcatagtttctc 406

OY 326 caggtttgaggggcccgagatgtggaaccccaacccgtgagctgagcggagagctgctgt 385

Db 407 caggcttcctatgatcagatgtggaaccccaacactgacactgagctgagagctgttat 466

OY 386 acctcaacgtgtggacacataccccccggcctacatccccccacctctcctcgtctgga 445

Db 467 atctaaatgtatgattccagcacctaaacaaaataatgcca--ctgtatgatatgga 523

OY 446 totatgggggtggttctacagtggggcctcctcctcttggagctgtaacgtggcgccttc 505

Db 524 ttatggtggtgttttcaaacctggaacatcatctttacatgtttatgatggcaagtctc 583

OY 506 tgggtacagcccgagagagactgtgctggtgtccatgaactaccgggtgggagccttggct 565

Db 584 tggctcgggttgaaagaggttatgtagtgccaatgaactataggtgggtggcctcaggat 643

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:25:56 ; Search time 6064.26 Seconds
(without alignments)
3839.256 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgagggcccccagtgctct.....aattgctcagcgccactga 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length	DB	ID
1	1291	74.3	2080	11	BC001541
2	764.2	44.3	990	10	BI489087
3	750.4	43.5	931	10	BI488594
4	698.6	40.5	780	10	BI753192
5	688.6	39.9	822	9	AI190022
6	638.8	37.0	752	10	BI917862
7	638.2	37.0	781	10	BG705869
8	623.8	36.2	786	10	BG703349
9	618.4	35.8	732	10	BG708331
10	615.4	35.7	738	10	BI667712
11	588.8	34.1	702	10	BI915019
12	560.2	32.5	589	9	AI654042
13	558	32.3	562	10	BE466391
14	555.2	32.2	689	10	BG395920
15	547	31.7	722	10	BE535653
16	538.8	31.2	555	9	AI769167
17	529.8	30.7	534	9	AI332425

C 18	521.6	30.2	577	9	AI802781	AI802781 wf18c08.x
C 19	494.6	28.7	549	9	AI143271	AI143271 qc04607.x
C 20	484.6	28.1	574	9	AI207928	AI207928 ap14a04.x
C 21	479.6	27.8	533	9	AA406567	AA406567 zv11a11.s
C 22	476.4	27.6	478	9	AI082805	AI082805 ox74d09.x
C 23	474	27.5	507	9	AW196326	AW196326 xm31h05.x
C 24	469.8	27.2	474	9	AI761618	AI761618 wg66f01.x
C 25	467.8	27.1	483	9	AI761619	AI761619 wg66f02.x
C 26	461	26.7	483	9	AI761609	AI761609 wg66e02.x
C 27	454.4	26.3	456	9	AI081892	AI081892 ox77b06.x
C 28	454.4	26.3	932	10	BF235425	BF235425 602025226
C 29	444.4	25.8	625	10	BI464924	BI464924 603207402
C 30	432.4	25.1	485	9	AA447461	AA447461 zw89g10.r
C 31	432	25.0	514	9	AA032228	AA032228 zf01f07.s
C 32	427.8	24.8	489	9	AI091215	AI091215 oc22g11.x
C 33	424.4	24.6	586	9	AA446656	AA446656 zw89g10.s
C 34	420.6	24.4	523	10	BF191608	BF191608 23941.MA
C 35	417	24.0	525	10	BM256199	BM256199 518509.MA
C 36	414.6	24.0	506	10	BE721978	BE721978 189843.MA
C 37	412.6	23.9	550	10	BE808194	BE808194 213342.MA
C 38	412.6	23.9	558	10	BF074683	BF074683 222160.MA
C 39	402.4	23.3	405	9	AW050645	AW050645 wz19f11.x
C 40	401.6	23.3	495	9	AA411703	AA411703 zv11a11.r
C 41	401	23.2	821	10	BE540575	BE540575 601067212
C 42	397.2	23.0	662	9	BB629080	BB629080 BB629080
C 43	396.4	23.0	556	9	AI332839	AI332839 qp96g08.x
C 44	382	22.1	504	10	BM286130	BM286130 526062.MA
C 45	377.2	21.9	396	10	H21132	H21132 yn65b07.r1

ALIGNMENTS

RESULT 1
BC001541
LOCUS BC001541 2080 bp mRNA linear HTC 31-JAN-2002
DEFINITION Homo sapiens, clone IMAGE:3453562, mRNA.
ACCESSION BC001541
VERSION BC001541.1 GI:14705895
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2080)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 4 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156
This clone has the following problem: no polyA-tail.
FEATURES
Location/Qualifiers
1..2080
/organism="Homo sapiens"

```
/db_xref="taxon:9606"
/clone="IMAGE:3453362"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
BASE COUNT      319 a      692 c      651 g      418 t
ORIGIN

Query Match      74.3%   Score 1281;   DB 11;   Length 2080;
Best Local Similarity 82.68;   Pred. No. 1e-234;
Matches 163;   Conservative 0;   Mismatches 0;   Indels 346;   Gaps 1;

Qy 1 atgagggccccgcagtgctgtctgcacacgccttccctggcttcccccactcttctctc 60
Db 98 ATGAGGCCCCGCGAGTGTCTGCTGCACACGGCTTCCCTGGCTTCCCNACTCTCTCTC 157

Qy 61 ctctctggctctctgggtgagggagtgagggtgagggcgaggagtgagagctgctg 120
Db 158 CTCTCTGGCTTCTTGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGCGAGAGCTGCTG 217

Qy 121 gtgacggtgctggggcgctgcgggcatctgcctgaagaccccgggggccctgtc 180
Db 218 GTGACGGTGGTGGGGCGGCTGCGGGGCTTCCCTGAAGACCCCGGGGGCCCTGTC 277

Qy 181 tctgtcttctgggcatcccttgcggagccacccatggagcccgctgcgtttctgcca 240
Db 278 TCTGTTTCTCTGGGATCCCTTTTGGGAGCCACCCATGGGACCCCTGCTTCTGCCA 337

Qy 241 ccggagagccagagccttggtcaagggtgtagagctacaaactccagagtgctgc 300
Db 338 CCGAGAGCCCAAGCAGCCTTGGTCAAGGGGTGGTAGACGTACAACCTTCCAGAGTGTCTGC 397

Qy 301 taccaatatgtggacacccatataccagggtttgagggcaccagagatgtgaaccccaac 360
Db 398 TACCAATATGTGGACACCCCTATACCCAGGTTTTGAGGGCACCGAGATGTGAACCCCAAC 457

Qy 361 cgtgagctgagcaggagctgctgtacctcaacgtgtggacacatacccccgccctaca 420
Db 458 CGTGAGCTGAGCGAGGACTGCTGTACTCAACGTGTGGACACCATACCCCGGSCCTACA 517

Qy 421 tccccacccctgtcctgtctgtagatctatgggggtggcttctacagtggggcctctcc 480
Db 518 TCCCCACCCCTGTCCTGCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCC 577

Qy 481 ttgacgtgtacagtgccgtcttctgttacagccagagagactgtctggtgtccatg 540
Db 578 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGCCCGAGAGGACTGTGTGGTGTCCATG 637

Qy 541 aactaccgggtggagagccttggcttccctgcccctgcggggagccgagagcccccgggc 600
Db 638 AACTACCGGGTGGAGGCTTTGGCTTCTGTGCCCCGCGGGAGCCGAGAGGCCCGCGGC 697

Qy 601 aatgtgggtctctgtagatcagaggtggccctgcagtgaggtgagagaaactgtgagacc 660
Db 698 AATGTGGGTCTCTTGGATCAAGAGCTGGCCCTGCAGTGGGTGCGAGAGAACTGTGGCAGCC 757

Qy 661 ttccgggggtaccgcacacagtgacgtgtttggggagagcgagcgagccgctcggtg 720
Db 758 TTCCGGGGGTGACCCACATCAGTACGACCTGTTTGGGGAGAGAGCGCGGAGCGCTCGGTG 817

Qy 721 ggcattgacattgtgtcccccgccagccggggcctgtttccacaggggcctgtgtagagc 780
Db 818 GGCATGACCTGCTGTGCCCCGCCAGCGCGGGGCTGTTCACAGGGGCGCTGCTGCAGAGC 877

Qy 781 ggtgcccccaatggagcctgggcaacggtgggcatgggagagggccctgcaggggcaag 840
Db 878 GGTGCCCCCAATGGAGCCTGGGCCACAGGTGGCATGGGATGGAGAGGCCCTGCGAGGGCCAGC 937

Qy 841 cagctggccccacactgtgggctgtctccacaggggcactggtgggaatgacacagagctg 900
Db 938 CAGCTGGCCCCACACTTGTGGGCTGTCTCCAGGGCGGCTCAGCAGTACGTTAGTCTGGA 2077
```

Qy	1635	cct 1637	
Db	2078	CCT 2080	
RESULT 2			
BI489087/c			
LOCUS			
DEFINITION			
603021122R1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 3',			
mRNA sequence.			
ACCESSION			
BI489087			
KEYWORDS			
EST.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
1 (bases 1 to 990)			
NIH-MGC http://mgc.nci.nih.gov/ .			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs@email.nih.gov			
Tissue Procurement: Life Technologies, Inc.			
cDNA Library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1AM11479 row: 1 column: 06			
High quality sequence stop: 803.			
Location/Qualifiers			
1. 990			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:5191829"			
/clone_lib="NIH_MGC_114"			
/lab_host="DH10B"			
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;			
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6			
male brains, age range 23-27 yo. Library is oligo-dT			
primed and directionally cloned {EcoRV site is destroyed			
upon cloning}. Average insert size 1.5 kb, insert size			
range 1-3 kb. Library is normalized and enriched for			
full-length clones and was constructed by C. Gruber			
(Invitrogen). Research Genetics tracking code 019. Note:			
this is a NIH_MGC Library."			
BASE COUNT			
174 a 316 c 318 g 182 t			
ORIGIN			
Query Match			
44.38; Score 764.2; DB 10; Length 990;			
Best Local Similarity			
96.00; Pred. No. 4.1e-136;			
Matches 869; Conservative 0; Mismatches 28; Indels 8; Gaps 8;			
Qy	733	ctgtcccccagccagcgggcgctgtccacagcgccgtgctgcagagcggtgcccccaat 792	
Db	898	CTCTGTTCCCGCCAAACGGGGACGTGTCCACAGGCCCTGCTGCAGACGGTGCCTCC	
Qy	793	ggaccctggggccacggtggcattgggagagcccgctgcagggccacgacgtgcccac 852	
Db	838	GGA-CCTGGGCCACGTGGGCCATGGAAAGCCCGTCGCA-GGCCACGACGTGGCCAC 781	
Qy	853	cttgggctgtctccacagcgccgacctggtgggaatgaca-cagagctggtagctgcct 911	
Db	780	CTTGT-GGCTGTCTCTCCA-GCGGCACCTGTTGGGAATGACACAGAGCTGTGTAGCTGCCT 723	
Qy	912	tcggacacgaccagcgaggtccctgggtgaaccagcaatggcacgtgctgcctcaagaag 971	
Db	722	TC-GNACAGCACGGCAGGTCTGTGGTGAACACCAAGATGGCAGCTGTCTCCTCANGAAG 664	
Qy	972	cgtcttcgggttctcctctcgtgcctgtggtagatggagacttctcagtgacacccaga 1031	

[illegible]

Plate: LLAM11479 row: 1 column: 06
High quality sequence stop: 771.
Location/Qualifiers

FEATURES
source

1. .931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5191829"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed): RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 152 a 294 c 306 g 179 t
ORIGIN

Query Match 43.5% Score 750.4; DB 10; Length 931;
Best Local Similarity 96.9%; Pred. No. 1.7e-133;
Matches 882; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

Qy 130 cgtggggcgccgtcgccggttcgctgaagacccccggggccctgtctctgttc 189
Db 1 CGTGGGGCGCGCTCGGGGGCATTCGCTGAAGACCCCGGGGGCCCTGTCTCTGCTTC 60
Qy 190 ctggggatcccccttgcggagccaccatgggaccccgctgccttctgcaccggagccc 249
Db 61 CTGGGCATCCCTTTTGGGAGCCACCATGGACCCCGCTGCTTCTGCCACCGAGCCC 120
Qy 250 aagcagccttgtaggggtgtagagcctacaacctccagagtgctgtctacaatat 309
Db 121 AAGCAGCCTTGCTCAGGGGTGTAAGCTACAACCTTCCAGAGTGCTGTACCAATAT 180
Qy 310 gtggacacccctataccagggttttggggccaccgagatgtggaaccccaacccgtgagctg 369
Db 181 GTGGACACCCCTATACCCAGGTTTTGAGGGCACCCGAGATGTGGAACCCCAACCGTGAGCTG 240
Qy 370 agcagagactcctgtacctcaacagtgtagacacataccccccggcctacatacccccaacc 429
Db 241 AGCGAGGACTGCCCTGTACCTCAAGCTGTGGACACCATACCCCGGCGCTACATCCCCCACC 300
Qy 430 cctgtcctcgtctggtatctatgggggtggctctacagtggggctcctccttg-gacgt 488
Db 301 CCTGCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGGCTCTCTCTGTGAGCT 360
Qy 489 gtacgatggccgctcttgggtacagggccgagagactgtgtgtgtccatgaactaccg 548
Db 361 GTACGATGGCGGCTTCTTGGTACAGGCGGAGAGGACTGTGTGGTGTCCATGAATAACCG 420
Qy 549 ggtggagcctttggcttcccttcggggagccgagagccgagagcccgccgaatgtggg 608
Db 421 GGTGGGAGCCCTTGGCTTCTGCGCCCTGCGGGGAGCGGAGAGCGCCCGGCAATGTGGG 480
Qy 609 tctcctgatcagaggtcgccctgcagtggtgcaggagacgtggcagccttcggggg 668
Db 481 TCTCTGGATCAGAGGCTGGCCCTGCAAGTGGGTGACAGAGAACGTGGCAGGCTTCGGGGG 540
Qy 669 tgacccacatcagtagcgtgtttggggagagcgcgggagccgctcgtggggcatgca 728
Db 541 TGACCCGACATCAGTACGCTGTGTTGGGAGAGCGCGGGAGCGCCCTCGGTGGGATGCA 600
Qy 729 cctgtgtccccccagccagccggg-cctgttccacagggccgctgtcagagc-ggtgccc 786
Db 601 CCTGCTGTCCCGCCAGCGGGGCGCTGTTCACAGAGGGCGCTGTGCAGAGCGGGTGC 660
Qy 787 cccaatggacctgggcaacggtggcatggagagagcccgctcaggggacac--gcagc 844
Db 661 CCCAATGGACCTTGGGCCACGGTGGGATGGAGAGGCCCTGTCGACAGGGCCACAGCAGCT 720

Qy 845 tggccacccttgggtgtgt-cttccaggcgccactggtgggaatgacacagagct-ggt 902
Db 721 TGGCCCACTTGTGGGTGTCCCTCCAGCGGCACTGGTGGGAATGACACAGAGCTGGGT 780
Qy 903 agcctgcc-ttcggacacgacc-aggcaggctcctggtgaaccacgaaatggcagctgctg 960
Db 781 AGCTGCTCTTTCGGACACGACCAAGCGCAGGTCTGTGTAACACAGAAATGGCAGCTGCTG 840
Qy 961 cctcaagaagcgtcttcctcgttctcctcgtcgtggtgtagatggagacttctcagtg 1020
Db 841 -CTCAAGNAACGCTCTCCGG-TCTCCTTCGTGCTGT-GTCCATTCGCACTTCTCTCAGG 897
Qy 1021 gacacccacg 1030
Db 898 AAACCCACAG 907
RESULT 4
BI753192 780 bp mRNA linear EST 25-SEP-2001
LOCUS 603026066F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196256 5',
DEFINITION mRNA sequence.
ACCESSION BI753192
VERSION BI753192.1 GI:15744770
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS NTH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 row: d column: 17
High quality sequence stop: 780.
FEATURES
source
1. .780
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196256"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed): RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 111 a 255 c 260 g 154 t
ORIGIN

Query Match 40.5% Score 698.6; DB 10; Length 780;
Best Local Similarity 98.4%; Pred. No. 1.3e-123;
Matches 717; Conservative 0; Mismatches 9; Indels 3; Gaps 1;
Qy 1 atgaggcccccgagtgctgtcgtcacacgcttccctcgtggtcccccactctctctc 60
Db 48 ATGAGGCCCCGCGAGTGCTGCTGCACACGCGCTCCCTGGCTTCCCCACTCTCTCTC 107

```
Qy 61 ctccctgctcctggtgagagagtggtgggctgagggccggagagatgacagagctgctg 120
Db 108 CTCTCTCTGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGATGCAGAGTGTCTG 167
Qy 121 gtgaggtgctgtggggccggctgctggggcattccctgaaagaccccccgggggcccctgtc 180
Db 168 GTGACGTGTGCTGGGGCCGGCTGCGGGCATTCGCCCTGAACACCCCGGGGGCCCTGTC 227
Qy 181 tctgttctctgggcatccctcttctggagagccaccatcagggaccctcgtctgttttgcca 240
Db 228 TCTGTTCTCTGGCATCCCTTTTCGGAGACCACCATGGGACCCCGCTGCTTCTTGCCA 287
Qy 241 ccgagagcccaagacagccttgctcaggggtgtgtagcgtacacactccagagtgctgc 300
Db 288 CCGGAGCCCAACACAGCTTGTTCAGGGGTGTAGACGCTTACAACTTCCAGAGTGTCTGC 347
Qy 301 taccatattgtgacacccctataccacaggttttggggccacagagatgtggaaccccaac 360
Db 348 TACCAATATGTGACACCCCTATACCAGGTTTGTAGGGCACCGAGATGTGGAACCCACAC 407
Qy 361 cgtgagctgagcagagactgctgtacctcaacgtgtgagacaccatacccccggcctaca 420
Db 408 COTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCTACA 467
Qy 421 tccccaccctctcctcgtctggtatctatgggggtggtcttctacagtgggcccctctcc 480
Db 468 TCCCCACCCCTCTCCTGCTCTGATGATATGGGGGTGGCTTCTACAGTGGGGCCCTCTCC 527
Qy 481 ttggacgtgtacagtgccgcctcttctgtacagggccgagagactgctggtgtccatg 540
Db 528 TTGAGCTGTACGATGGCGCTTCTTGTGTACAGCCGAGAGACTGTGCTGTGTCTCATG 587
Qy 541 aactaccgggtggggcccttctgctcctctcctgcccctgcccgggagccgagagccccgggc 600
Db 588 AACTACCGGGTGGGACCTTTTGGCTTCTGTGGCCCTGCCGGGAGCCGAGAGGCCCGGGGC 647
Qy 601 aatgtggtctcctgacagagctgcccctgcagtgagtgagtgagagagagagagagagagc 660
Db 648 AATGTGGGTCTCCTGGATCAGAGGCTTGGCCCTGCGAGTGGGTGCAGAGAACTGTGCAGCC 707
Qy 661 ttccgggggtgacccacacatcagtcagcgtgttggggagagcgcggggagccctcctggtg 720
Db 708 TTCCGGGGGTGACCCGACATCATGACGCTGT--GGGAGAGCGCGGAGCGCGCTCGGTG 764
Qy 721 ggcattgcac 729
Db 765 GGCATGCC 773

RESULT 5
A1190022/c
LOCUS
DEFINITION
IMAGE:1731502.3' similar to gb:M5040 ACETYLCHOLINESTERASE
PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 822)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LUNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1019 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
```

FEATURES
source

High quality sequence stop: 469.

Location/Qualifiers

1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1731502"
/clone_lib="Soares_fetal_heart_NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site:1: Not 1; Site:2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCCGACATCTTTTCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaudo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHH19W."

BASE COUNT 151 a 271 c 245 g 150 t 5 others
ORIGIN

Query Match 39.9%; Score 688.6; DB 9; Length 822;
Best Local Similarity 93.8%; Pred. No. 1.1e-121;
Matches 769; Conservative 0; Mismatches 44; Indels 7; Gaps 5;

Qy 833 gggccacagcagctggcccacacctgtggtgtgtctctccacggcgagcactgg--tgggaatga 890
Db 820 GCCCAACAGAGTGGCCCCCCTTGGTGGGTGTCTCTCAGCGGGCATGTGGTGGGAATGA 761
Qy 891 cacagagctgtagctgtccttcgacacagcagcgca-gtctcctggtggaaccacgaat 949
Db 760 CACAAAAGTGTGTAGTGTGCTTCGGACACACACAGGGCAGGGTCTTTGGTGAACACGAATG 701
Qy 950 ggcacgtgtcgtcctcaag-aagcgtcttcctcggttctec--ttcgtgcctgtggtagatg 1006
Db 700 GGCACGTGTGTCTCAAGAAAGCGTCTTCGGGTCTCCTTCNGTCCCTGTGGTAGATG 641
Qy 1007 gagacttctcagtgacacccacagag-gcctcatcaacgcgggagaaacttccacggcctg 1065
Db 640 NAGACTCTCAAGTGACACGCCAGAGGCCCTCATCAACGCGGGAGACTTCACGGGCGCTG 581
Qy 1066 caggtgtgtgtgtgtgaagatgaggtcgtatatttctgttttctacggggccccc 1125
Db 580 CAGGTGTGTGTGGTGTGTGAAGATGAGGGCTCGTATNTTCTGTTTACGGGGCCCCA 521
Qy 1126 ggtctcagcaagacaacagagctctctcatcagccggcgagttcctgcccgggtgcgg 1185
Db 520 GGTTCAGCAAAAGACAACAGTCTCTCATCAGCGGGCGAGTTCCTTGCGCGGGTGCAG 461
Qy 1186 gtccgggttccccaggtaagtacacctggcagccagagctgtgttctctgattacacagac 1245
Db 460 GTCCGGGTTCCCCAGGTAAGTGAACCTGGCAGCGAGGCTGTGTCTGCTGCTGCTGCTGCTG 401
Qy 1246 tggctgcatcccgaggaacccgcagcctgagggagccctcctgagcagatgtgtgtggcgac 1305
Db 400 TGGCTGCATCCCGAGGACCCCGCACGCCCTGAGGGAGGCCCTGAGCGATGTGTGTGGCGGAC 341
Qy 1306 cacaatgtcgtgtgcccctgtgcccagctggtgtggcgagctggctcccagaggtgcccgg 1365
Db 340 CACAATGTGTGTCNCCGTGGCCCCAGCTGGTGGCGAGTGGCTGCCAGGGGTGCCCGG 281
Qy 1366 gtctacgctacgtctttgaacacccgtgttccacgctctcctgtgcccctgtgtgtgggg 1425
Db 280 GTCTACGCCCTACGCTCTTTGAACACCGTGTCTCCACGCTCTCCTGGGCCCTGTGTGGATGGG 221
Qy 1426 gtgccccacgctacagatcagttcatcttggatccccctcctgacccccctcctcgaac 1485
Db 220 GTGCCCTCACGGCTACGAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAAC 161

Qy	1486	tacacggcagaggagaaatcttcgccacgcgactgatgcatactatggggcaatttggcc	1545
Db	160	TACACGGCAGAGGAGAAAATCTTCGCCACGAGCTGATGCGATACTGGGCCCAACTTTGCC	101
Qy	1546	cgcacagggatcccaatgaccccgaccccgagaccccaagcccccacacaatggccccctacacg	1605
Db	100	CGCACAGGGATCCCATGATGACCCCGAGACCCCAAGGCCCCACCAATGGCCCCCGTACACG	41
Qy	1606	gcgggggctcagcagtcacgtactgttagcttgacactgcggccgc	1645
Db	40	CGCGGGGCTCAGCAGTACGTTAGTCTGAGACCTCGCGCGCGC	1
RESULT	6		
BI917862	BI917862	752 bp	mrna
LOCUS	603183996F1 NIH_MGC_121 Homo sapiens	cdna	clone IMAGE:5248006 5',
DEFINITION	752 bp	mrna	linear
EST	EST 17-OC7-2001		
ACCESSION	BI917862		
VERSION	BI917862.1	GI:16199790	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	N1 (bases 1 to 752)		
JOURNAL	NCBI-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabps-femail.nih.gov Tissue Procurement: Life Technologies, Inc. cdna Library Preparation: Life Technologies, Inc. cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L14M11625 row: p column: 23 High quality sequence stop: 752. Location/Qualifiers		
FEATURES	source		
	1..752		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5248006"		
	/clone_lib="NIH_MGC_121"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed): RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."		
BASE COUNT	106 a	248 c	251 g
ORIGIN			147 t
Query Match	37.08;	Score 638.8;	DB 10; Length 752;
Best Local Similarity	98.7%;	Pred. No. 3.4e-112;	
Matches	665; Conservative	0; Mismatches	7; Indels
			2; Gaps
Qy	1	atgaggccccgcagtgctctgtcacacgccttcctcgcttccctccactctctctc	60
Db	80	ATGAGCCCCCGCAGTGCTCTGCTGCACAGCGCTTCCCTGGCTTCCCACTCTCTCTC	139
Qy	61	ctctcttggtctcttgggttgaggagtgaggggcgaggagatgcagagctgctg	120
Db	140	CTCCTCTGGCTTCCTGGGTGGAGAGTGGGGGCTTCAGGCGCGGAGATGCAGAGCTGCTG	199

```
/db_xref="taxon:9606"  
/clone_image="4792046"  
/clone_lib="NIH_MGC_96"  
/tissue_type="hypothalamus"  
/lab_host="DH10B"  
/note="Organ: brain; Vector: pBluescriptR (modified  
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag  
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.3 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIMH/NHGRI, National  
Institutes of Health). Note: this is a NIH_MGC Library."  
BASE COUNT      110 a 260 c 262 g 149 t  
ORIGIN  
  
Query Match      37.0%; Score 638.2; DB 10; Length 781;  
Best Local Similarity 99.3%; Pred. No. 4.4e-112;  
Matches 662; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
  
QY 1 atgagccccgcagtgctgtgcacacgcttccctggcttccccactctctctc 60  
DB 113 ATGAGCCCCCGCAGTGTCTGTGCACAGCCTTCCCTGGCTTCCCACTCTCTCTC 172  
  
QY 61 ctctctgtgctcctgggtgagagtggtgggctgagggccggagagtcagagtgctg 120  
DB 173 CTCTCTTGGCTCTGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGACAGTGTCTG 232  
  
QY 121 gtgacggtgctggggccgctcgagcattcgctgaagaccccggggcccctgtc 180  
DB 233 GTGACGGTGTGGGGCCGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGG 292  
  
QY 181 ttctgttctctgggcatccctttgcggagccacccatgggaccccgctgcttctgcca 240  
DB 293 TCTGCTTCTTGGGCATCCCTTTGGGGAGCCACCACCATACCCCGGCCCTACA 352  
  
QY 241 ccggagcccaagcagccttgggtcaggggtggtagcgtacacattccagagtgctgc 300  
DB 353 CCGGAGCCCAAGCAGCCTTGGTTCAGGGGTGGTACGCTTACAACTTCCAGAGTGTCTGC 412  
  
QY 301 taccatattgagacaccctataccaggttttgagccacgagatgtgaacccaac 360  
DB 413 TACCATATGTGGACACCCTATACCAGGTTTGGGGCACCAGATGTGGAACCCCAAC 472  
  
QY 361 cgtgagctgagcagactgctgtacctcaacgctgtgagaccatacccccggcctaca 420  
DB 473 CGTGAGCTGAGGAGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA 532  
  
QY 421 tccccacccctgtcctgctgtggtatctatggtgggtggtcttcaagtggtgctctcc 480  
DB 533 TCCCCCA-CCCTGTCTGCTCTGATCTATGCGGTGGCTTCTACAGTGGGGCTCTCTCC 591  
  
QY 481 ttggagctgtacgagtcgctcttctgtacagccgagagacgtgctgtgtccatg 540  
DB 592 TTGGACGTGTACGATGGCCGCTCTTGGTACAGCCGAGAGACGTGTGCTGTGTCCATG 651  
  
QY 541 aactaccgggtggagcctttgcttctcctggcctggcgggagccgagagccccgggc 600  
DB 652 AACTACCGGGTGGAGCCTTTTGGCTTCTTGGCCCTGCCGGGAGCGGAGAGCCCCGGGC 711  
  
QY 601 aatgtgggtctcctggaatgaagggctgcccctgagtggtgggtgcaggagaacgtgagcc 660  
DB 712 AATGTGGGTCTCTGGATCAGAGGCTGG-CCGTGACGTGGGTGCAGGAGAACCTGGCAGCA 770  
  
QY 661 ttcgggg 667  
DB 771 TCGGGGG 777
```

RESULT 8
BG703349

```
LOCUS      BG703349      786 bp      mRNA      linear      EST 07-MAY-2001  
DEFINITION 602685108F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4817881 5',  
mRNA sequence.  
ACCESSION  BG703349  
VERSION    BG703349.1  GI:13975591  
KEYWORDS   EST.  
SOURCE      human.  
ORGANISM   Homo sapiens  
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  1 (bases 1 to 786)  
            NIH-MGC http://mgs.nci.nih.gov/  
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE       Unpublished (1999)  
JOURNAL     Contact: Robert Strausberg, Ph.D.  
COMMENT     Email: cgapbs-remail.nih.gov  
            Tissue procurement: Miklos Palkovits, M.D., Ph.D.  
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
            Toshiyuki and Piero Carninci (RIKEN)  
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
            DNA Sequencing by: Incyte Genomics, Inc.  
            Clone distribution: MGC clone distribution information can be  
            found through the I.M.A.G.E. Consortium/LLNL at:  
            http://image.llnl.gov  
            Plate: LLNL10718 row: o column: 02  
            High quality sequence stop: 731.  
            Location/Qualifiers  
                1..786  
                    /organism="Homo sapiens"  
                    /db_xref="taxon:9606"  
                    /clone_image="4817881"  
                    /clone_lib="NIH_MGC_95"  
                    /tissue_type="hippocampus"  
                    /lab_host="DH10B"  
                    /note="Organ: brain; Vector: pBluescriptR (modified  
                    pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag  
                    ); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
                    size-selected for average insert size 2.5 kb and  
                    normalized to ROT 5. This is a primary library enriched  
                    for full-length clones and constructed using the  
                    Cap-trapper method (Carninci, in preparation). Library  
                    constructed by M. Brownstein (NIMH/NHGRI, National  
                    Institutes of Health). Note: this is a NIH_MGC Library."  
BASE COUNT  110 a 263 c 261 g 152 t  
ORIGIN
```

```
Query Match      36.2%; Score 623.8; DB 10; Length 786;  
Best Local Similarity 97.2%; Pred. No. 2.5e-109;  
Matches 656; Conservative 0; Mismatches 17; Indels 2; Gaps 2;  
  
QY 1 atgagccccgcagtgctgtgcacacgcttccctggcttccccactctctcctc 60  
DB 113 ATGAGCCCCCGCAGTGTCTGTGCACACGCTTCCCTGGCTTCCCACTCTCTCTC 172  
  
QY 61 ctctctgtgctcctgggtgagagtggtgggctgagggccggagagtcagagtgctg 120  
DB 173 CTCTCTTGGCTCTGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGACAGTGTCTG 232  
  
QY 121 gtgacggtgctggggccgctcgagcattcgctgaagaccccggggcccctgtc 180  
DB 233 GTGACGGTGTGGGGCCGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGGGGGCTTGG 292  
  
QY 181 ttctgttctctgggcatccctttgcggagccacccatgggaccccgctgcttctgcca 240  
DB 293 TCTGCTTCTTGGGCATCCCTTTGGGGAGCCACCACCATACCCCGGCCCTACA 352  
  
QY 241 ccggagcccaagcagccttgggtcaggggtggtagcgtacacattccagagtgctgc 300  
DB 353 CCGGAGCCCAAGCAGCCTTGGTTCAGGGGTGGTACGCTTACAACTTCCAGAGTGTCTGC 412  
  
QY 301 taccatattgagacaccctataccaggttttgaggggacccagagatgtgaacccaac 360  
DB 413 TACCATATGTGGACACCCTATACCAGGTTTGGGGCACCAGATGTGGAACCCCAAC 472  
  
QY 361 cgtgagctgagcagactgctgtacctcaacgctgtgagaccatacccccggcctaca 420  
DB 473 CGTGAGCTGAGGAGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA 532  
  
QY 421 tccccacccctgtcctgctgtggtatctatggtgggtggtcttcaagtggtgctctcc 480  
DB 533 TCCCCCA-CCCTGTCTGCTCTGATCTATGCGGTGGCTTCTACAGTGGGGCTCTCTCC 591  
  
QY 481 ttggagctgtacgagtcgctcttctgtacagccgagagacgtgctgtgtccatg 540  
DB 592 TTGGACGTGTACGATGGCCGCTCTTGGTACAGCCGAGAGACGTGTGCTGTGTCCATG 651  
  
QY 541 aactaccgggtggagcctttgcttctcctggcctggcgggagccgagagccccgggc 600  
DB 652 AACTACCGGGTGGAGCCTTTTGGCTTCTTGGCCCTGCCGGGAGCGGAGAGCCCCGGGC 711  
  
QY 601 aatgtgggtctcctggaatgaagggctgcccctgagtggtgggtgcaggagaacgtgagcc 660  
DB 712 AATGTGGGTCTCTGGATCAGAGGCTGG-CCGTGACGTGGGTGCAGGAGAACCTGGCAGCA 770  
  
QY 661 ttcgggg 667  
DB 771 TCGGGGG 777
```

```

Db 413 TACCAATATGTGGACACCCCTATACCCAGGTTTGTGAGGGCACCGAGATGTGAACCCCAAC 472
QY 361 cgtgagctgagcaggactgctgtacctcaacgtgtggacacacataccccccggcctaca 420
Db 473 CGTGAGCTGAGCGAGGAGTGCCTGTACTCAACGTGTGGACACCATACCCCGGGCTACA 532
QY 421 tccccacacctgctctcgtcgtgatctatgagggtgaggggtctctacagtggggctctcc 480
Db 533 TCCCCCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592
QY 481 ttggacgtgtacatgagccgctctctgtgtacagccgagagagactgtgtgtgtccatg 540
Db 593 TTGGACGTGTACATGAGCCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652
QY 541 aactacgggtgagagcctttggg-ttccctggccctccggggagcagagagcgggggg 599
Db 653 AACTACCGGGTGGAGAGCTTTGGCTTTCCCTGGCCCTGCCGGGAGCGGAGAGG-CCCGGG 711
QY 600 caatgtgggtctcctgagatcagagggctgcccctgcaagtgggtgcagagagaactggcagc 659
Db 712 CAATGTGGGTCTCTGATCACAGCTGGCTGCAGGTGGTGCAGGTGGTGCAGGTGGTGCAG 771
QY 660 ctccgggggtgaccc 674
Db 772 CTTGGGGGTGACCC 786

RESULT 9
Bg708331 732 bp mRNA linear EST 07-MAY-2001
LOCUS 602672268F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795048 5',
DEFINITION mRNA sequence.
ACCESSION Bg708331
VERSION Bg708331.1 GI:13985566
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 732)
AUTHORS NIH-MGC http://mgi.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0677 row: g column: 17
High quality sequence stop: 732.
Location/Qualifiers
1. .732
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4795048"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to Rf 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
101 a 250 c 239 g 142 t
BASE COUNT

```

```

ORIGIN
Query Match 35.8%; Score 618.4; DB 10; Length 732;
Best Local Similarity 99.8%; Pred. No. 2.6e-108;
Matches 619; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctgtgcacacgcttccctccctggcttcccccactctctctc 60
Db 113 ATGAGGGCCCCGCGAGTGTCTGTGCACACGCTTCCCTGGCTTCCCCACCTCTTCTCTC 172
QY 61 ctctctcgtctcctgggtgagagagtgaggggctgagggccgggagagatgagagctctg 120
Db 173 CTCTCTGTGGTCTGTGGTGTGGAGAGTGGGGCTGTAGGGCCGGAGAGTGCAGAGCTCTG 232
QY 121 gtacagtgctggtggggcgctgctgagggcattcgctcgaagaccccgggggccctctc 180
Db 233 GTGACGGTGTGGGGGGCGGCTGTGGGGGATTCGCTGTGAAGACCCCGGGGGCCCTGTC 292
QY 181 tctgcttctcgtgggcatccctcttgcggagccacccatggaccccgctcttcttgc 240
Db 293 TCTGCTTTCTGTGGCATCCCTTTTGTGGAGCACACCATGGACCCCGCTTCGCTTCTGCCA 352
QY 241 ccggagccccagcagccttggtcaggggtggtgagacactcaacacctccagagtgctgc 300
Db 353 CCGGAGCCCCAAGCAGCCTTGTGTGAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 412
QY 301 taccatattgtgacacccctataccacaggttttgagggcaccgagatgtggaaccccaac 360
Db 413 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGGACCGAGATGTGAACCCCAAC 472
QY 361 cgtgagctgagcaggagactgctgtacctcaacgtgtggacacacatacccccgccctaca 420
Db 473 CGTGAGCTGAGCGAGGAGTGCCTGTACTCAACGTGTGACACCATACCCCGGCTTACA 532
QY 421 tccccacacctgctcctgctgctgagatctatgagggtggtctctacagtggggctctcc 480
Db 533 TCCTCCCGAGCCCTGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 592
QY 481 ttggacgtgtacatgagccgctcttctgtgtacagccgagagagactgtgtgtgtccatg 540
Db 593 TTGGACGTGTACATGAGCCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652
QY 541 aactacgggtgagagccttctgctcctggccctgcgggggagcagagagggccgggc 600
Db 653 AACTACCGGGTGGAGAGCTTTGGCTTCTGTGGCTGTGGGGGAGCGGAGAGGCCCCGGGC 712
QY 601 aatgtgggtctcctggatca 620
Db 713 AATGTGGGTCTCTGTGGATCA 732

RESULT 10
Bg708712 738 bp mRNA linear EST 12-SEP-2001
LOCUS 603293067F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312273 5',
DEFINITION mRNA sequence.
ACCESSION Bg708712
VERSION Bg708712.1 GI:15581945
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 738)
AUTHORS NIH-MGC http://mgi.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)

```


CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11791 row: f column: 18
High quality sequence stop: 736.
Location/Qualifiers
1. 738

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312273"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0F 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 107 a 231 c 248 g 152 t

Query Match 35.7%; Score 615.4; DB 10; Length 738;
Best Local Similarity 97.6%; Pred. No. 9.9e-108;
Matches 657; Conservative 0; Mismatches 11; Indels 5; Gaps 3;

Qy 1 atgagggcccgagtgctgtgcacacgcgttccttgcgttccccactctctctc 60
Db 62 ATGAGGCCCCGACGTGCTGTGCACACGCTTCCCTGGCTCCCACTCCTTCCTC 121
Qy 61 ctctctgtctccctgggtggagagtgagggctgagggccgggagatgagctgtg 120
Db 122 CTGCTGTGGCTTCCGTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGATGCGAGAGTGTG 181
Qy 121 gtacaggtgcgtggggccggtgcggggcattgcctgaagaccccgggggccctgtc 180
Db 182 GTGAGGTTGCTGGGGGCGGCTGGGGGGCATTCGCTTGAAGACCCCGGGGCGCTGTC 241
Qy 181 tctgcttctcgggcatccctttgcggagacacccatggagaccctgtctgttttgc 240
Db 242 TCTGCTTTCTTGGGCATCCCTTTTGGGAGCACCCCATGGGACCCGCTGCTTTCGCA 301
Qy 241 ccgagcccaagcagcctgtgcaggggtgtgacgctacaacacctccagagtgctgc 300
Db 302 CCGGAGCCCAAGCAGCCTTGTGTAGGGGTGGTGTAGCGCTTACAACCTTCCAGAGTGTGCG 361
Qy 301 taccataatgtggacacctataccaggttttggggccacgagatgtgaaccccaac 360
Db 362 TACCAATATGTGGACACCTTATACCAGTATTGAGGGCACCAGAGATGTGGAACCCCAAC 421
Qy 361 cgtgagctgagcaggagactgctgtacctcaacgtgtggacacacataccccggcctaca 420
Db 422 CGTGAGCTGTAGCGAGGAGTGGCTGTGTACCTCAACGTGTGGACACCATACCCCGGCGCTACA 481
Qy 421 tccccaccctgtcctcgtctgtgatctatgggggtgtcttctacagtggggccctctcc 480
Db 482 TCCCCACCCCTGTCTATCGTGTGGATCATGGGGGTGGCTTCTACAGTGGGGCCCTCCCTCC 541
Qy 481 ttggagctgtacgagtcgctctctgttcagagccgagagactgtgc-tggtgtccat 539
Db 542 TTGGACGTGTACGATGGCGCGCTTCTTGTGTACAGCCGAGAGAGACTGTCTTGGTCTCCAT 601
Qy 540 gaactaccgggtggagagcc-----ttgggttctcctggccctgcggggagcccgaga-ggccc 595
Db 602 GAACGTACCGGGTGGGAGAGCCCTTTTGGCTTCTGTGGGCCCTGCGGGGGAGCCGAGAGGGCCC 661
Qy 596 cgggcaatgtgggtctctcgtgatcacagaggtggtccctcagtggtggcgaggaacagtgg 655

Db 662 CGGCAAGTGTGGGTCTCTGTGATCAGAGCTGGCCCTCAGTGGGTGAGGGAACGTTGG 721
Qy 656 cagccttcggggg 668
Db 722 CAGCCTTCGGGGG 734

RESULT 11
BI915019 702 bp mRNA linear EST 16-OCT-2001
LOCUS 603177204F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241624 5',
DEFINITION mRNA sequence.
ACCESSION BI915019
VERSION BI915019.1 GI:16179103
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 702)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11609 row: g column: 01
High quality sequence stop: 702.

FEATURES
source

1. 702
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5241624"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pcMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1-7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
BASE COUNT 99 a 240 c 226 g 137 t
ORIGIN

Query Match 34.1%; Score 588.8; DB 10; Length 702;
Best Local Similarity 99.4%; Pred. No. 1.2e-102;
Matches 612; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 atgagggcccgagtgctgtgcacacgcgttcctcgttccccactctctctc 60
Db 87 ATGAGGCCCCGACGTGCTGTGCACACGCTTCCCTGGCTCCCACTCCTTCCTC 146
Qy 61 ctctctgtcctcgttggtggagagtgagggctgagggccggagagatcacagctgtgc 120
Db 147 CTCCCTCTGGCTCCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGACAGCTGCTG 206
Qy 121 gtgacggtgcgtggggccggtgcggggcattcgcctgaagaccccgggggccctgtgc 180
Db 207 GTGACGGTGTGCTGGGGGCGGCTGCGGGGCATTTCGCTTGAAGACCCCGGGGCGCTGTC 266
Qy 181 tctgcttctcgtgggcatccctcttgcggagacacccatgggaccccgctcgtcttctgc 240

```
Db 267 TCAGCTTTCTGGGATCCCTTTGGGAGCCACCCATGGAGCCCGTCTGCTTTCTGCCA 326
QY 241 ccgagacccaagcagccttggtcaggggtgtagacgctacaacacctccagagtgctgc 300
Db 327 CCGGAGCCCAAGCAGCCTTGGTCAGGGTGGTAGACGCTACACACCTCCAGAGTGTCTGC 386
QY 301 taccatatgtgacacccctataccacaggtttgagggcaccagagatgtagaaccccaac 360
Db 387 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCAGAGATGTGGAACCCCAAC 446
QY 361 cgtgagctgagcagagactgctgtacctcaacgtgtggacaccatacccccgacctaca 420
Db 447 CGTGAGCTGAGCGAGGAGCTGCTGTACCTCAACGTGTGGACACCATACCCCGGACTACA 506
QY 421 tccccacacctctctcgtctgtgattctatggtggtctctacagtggggcctctcc 480
Db 507 TCCCCCAACCCCTGCTCTGCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCC 566
QY 481 ttggacgtgtacagatggcgcctctctgtgtacagggccag-aggaactgtgctggtgccat 539
Db 567 TTGGAGCTGTACGATGGCGCTCTTGTGTACAGCCGAGCAGGACTGTGCTGTGTCCAT 626
QY 540 gaactaccgggtggagccttggctctctgctcctgcccctgcccgggagccgagagggcccc-99 598
Db 627 GAACCTACCGGGTGGAGCCCTTGGCTTCTGTGGCCCTGCGCGGGAGCCGAGAGCCCGGG 686
QY 599 gcaatgtgggtctcct 614
Db 687 GCAATGTGGGTCTCCT 702

RESULT 12
AI654042/c
LOCUS
DEFINITION
AI654042 589 bp mRNA linear EST 07-MAR-2000
similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION
VERSION
AI654042.1 GI:4738021
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 589)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs@remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 850 Std Error: 0.00
Seq primer: -400p from gibco
High quality sequence source: 461.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2283550"
/clone_lib="NCI-CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site:1: Not 1; Site:2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_Kid3 was
```

```
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo.
BASE COUNT 106 a 193 c 186 g 101 t 3 others
ORIGIN
Query Match 32.5%; Score 560.2; DB 9; Length 589;
Best Local Similarity 99.3%; Pred. No. 3.2e-97;
Matches 562; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1055 tccacggcctcaggtgctggtggtgagatgagggctctgtttctgtgttt 1114
Db 566 TCACCGGCTTCAGAGTGTGGTGGGTGTGTGANGATGAGGGCTGTATTTCTGGTTT 507
QY 1115 acggggcccccaggttcagcaaaagacacagagtcctcatcagccgggcccagttcctgg 1174
Db 506 ACGGGGCCCCAGGCTTCAGAAAGACACAGAGTCTCTATCACCCGGCGGAGTTCCTGG 447
QY 1175 ccgggtgctgggtgctgggttccccaggtaagtacacctgagccgaggtggtgctctgc 1234
Db 446 CCGGGGTGGGGGTTCGGGGTTCGCCAGGTAAAGTACCTGGCAGCCGAGGCTGTGCTCTGC 387
QY 1235 attacacagactggtgctgctcccgagagcccgacgctgagggagccctgagcgatg 1294
Db 386 ATTACACAGACTGGCTGCATCCCGAGGACCCGACGCTTGAGGAGGCCCTGAGCGATG 327
QY 1295 tgggtgggacacacaaatgtctgtgctcccggtggccagctggtgggagctggtgccc 1354
Db 326 TGGTGGCGGACCAACAATGTCTGTGCNCGTGGCCAGCTGGCTGGCGGACGTGGCTGCC 267
QY 1355 aggggtcccggtctacgctacgtcttttgaaacacgctctccacgctctctctggcccc 1414
Db 266 AGGGTCCCGGGGTCTACGGCTTCTTTGAACCCCGTCTTCCACGCTCTCTCTGGCCCC 207
QY 1415 tgtggtatgggggtggcccccaggtctgctgagatgagttctcttggagctccctcgacc 1474
Db 206 TGTGGATGGGGTGGCCCCACGGCTACGAGATCGATGTTTCATCTTGGGATCCCGCTGGACC 147
QY 1475 cctctcgaaactacacggcagagggagaaatcttcccagcgactgactgctggg 1534
Db 146 CTTCTCGAAACTACACGCGCAGAGGAGAAAATCTTCGCCAGCGACTGATCGGATCTGGG 87
QY 1535 ccaactttcccgccacaggggatcccaatgagcccgagaccccaaggccccacaatggc 1594
Db 86 CCAACTTTGCCCGCACAGGGGATCCCAATGAGCCCGAGACCCCAAGGCCCAATATGGC 27
QY 1595 ccccgctacacggcggtggtcctcagcag 1620
Db 26 CCCCCTACACGGGGGGGCTCAGCAG 1
RESULT 13
BE466391/c
LOCUS
DEFINITION
BE466391 562 bp mRNA linear EST 27-JUL-2000
similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN
);contains element MSRI repetitive element ;, mRNA sequence.
ACCESSION
VERSION
BE466391
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 562)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE
```

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 469.
Location/Qualifiers

FEATURES
source

1. .562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3208591"
/clone_lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: p7R73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
88 circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
104 a 184 c 173 g 97 t 4 others

BASE COUNT
ORIGIN

Query Match 32.3%; Score 558; DB 10; Length 562;
Best Local Similarity 99.3%; Pred. No. 8.4e-97;
Matches 558; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1066 caggtgtctggtggtgaagatgagggctgtctgtattttctgttaccggggcccca 1125
|||||
Db 562 CAGGTGCTGTGTGTGAGGATGAGGCTCGTATTTTCTGGTNTACGGGGCCNCA 503
|||||
Qy 1126 ggccttcagaaagacaagctctctatcagccggccgagttcctgcccgggtgcgg 1185
|||||
Db 502 GGCCTCAGCANAGACAAGAGTCTCTATCAGCCGGCCGAGTTCTGTGCCGGGTGGCG 443
|||||
Qy 1186 gtccgggttccccaggtaagtgcctgagcagcagggctgtggtcctgcattacacagac 1245
|||||
Db 442 GTCCGGGTTCNCAGGTNAGTGACTGCGACCCGAGGCTGTGTGCTGCATTACACAGC 383
|||||
Qy 1246 tggctgcattcccgaggaccggccgctgagggagggccctgagcgtgtggtggcgac 1305
|||||
Db 382 TGGCTGCATCCGAGGACCCGGCAGCTGAGGGAGGCCCTGAGCGATGTGTGGCGCAG 323
|||||
Qy 1306 cacaatgtctgtgcccgcgtgagccagctgctggtggcgagctgctgcccaggggtgcgg 1365
|||||
Db 322 CACAATGTCGTGTGCCCGTGGCCAGCTGTGCTGGCGACTGGTGTGCCAGGGGTGCCGG 263
|||||
Qy 1366 gtctacgcctacgtcttttgaacacgctgtctccacgctctctcgtgcccctgtgattggg 1425
|||||
Db 262 GTCTACGGCTACGTCCTTTGAACACGGTCTTCACAGCTCTCTGCCCCCTGTGATGGGG 203
|||||
Qy 1426 gtgccccacggctacgagatcagttcatctttgggatccccctggagccccctctctgaaac 1485
|||||
Db 202 GTGCCCCACGGCTACGAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAAC 143
|||||
Qy 1486 tacacggcagaggagaaaattctgcaccagcagactgatacgatactggggcacaactttgcc 1545
|||||
Db 142 TACACGGCAGAGAGAAAATCTTCCGCCAGGCGACTGATGCGATACTGGGGCCAACTTTGCC 83
|||||
Qy 1546 cgcacagggggtatcccaatgagcccgagaccccaaggcccaatggccccctacacg 1605
|||||

Db 82 CGCACAGGGGATCCCATGAGCCCCGAGACCCCAAGGCCCCACATGCCCCCGTACAGC 23

Qy 1606 9cggggggtcgcagcagtagcttta 1627
|||||
Db 22 CGCGGGGCTCAGCAGTACGTTA 1

RESULT 14

BG395920 689 bp mRNA linear EST 12-MAR-2001
LOCUS 60245824F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580980 5',
DEFINITION mRNA sequence.
ACCESSION BG395920
VERSION BG395920.1 GI:13289368
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM1300 row: p column: 05
High quality sequence stop: 685.
Location/Qualifiers

FEATURES
source

1. .689
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4580980"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
94 a 234 c 223 g 138 t

BASE COUNT
ORIGIN

Query Match 32.2%; Score 555.2; DB 10; Length 689;
Best Local Similarity 98.1%; Pred. No. 3e-96;
Matches 604; Conservative 0; Mismatches 8; Indels 4; Gaps 4;

Qy 1 atgagggcccgagctgtctgtcacacgctctccctggctcccccactctctctc 60
|||||
Db 77 ATGAGGCCCGCAGTGTCTGCTGCACAGCGCTTCCCTGGCTTCCCACTCCTCTCTC 136
|||||
Qy 61 ctctctgtgctctctgggtgagggagtgagggtgagggccggagatgcagagctgtg 120
|||||
Db 137 CTCTCTGTGGCTCCTGGGTGGAGAGTGGGGGCTGAGGGCGGGAGAGTGCACAGCTGTG 196
|||||
Qy 121 gtgacgggtgctggggccggtcggtcggtcggtcggtcggtcggtcggtcggtcg 180
|||||
Db 197 GTGACGGGTGCTGGGGGCGGGCTGCGGGGCATTTCGCCCTGAAGACCCCGGGGCCCTGTC 256
|||||
Qy 181 tctgttttctctgggcatcccttttcgagagccacccatggagcccgctgcttctgcca 240
|||||
Db 257 TCTGCTTTCTTGGGGATCCCTTTTGGGAGGCCACCCATGGGACCGCTTCTTGCCA 316
|||||

QY 241 cggagcccaagcagccttggtgaggggtggtgagcgtacaccccttcagagtgctgc 300
 |||||
 Db 317 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACACCTTCACAGAGTGTCTGC 376
 QY 301 taccataatgtgagacacccataaccagattttgagggcaccagagatggaaccccaac 360
 |||||
 Db 377 TACCAATATGTGACACCCATATACCCAGG-TTGGAGGGACCCGAGATGTGAACCCCAAC 435
 QY 361 cgtgagctgagcagagactgctgtacctcaacgctgtgacacccaacccccggcctaca 420
 |||||
 Db 436 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGACACCATACCCCGGGCTACA 495
 QY 421 tcccccaaccctgtctctgctgagatctatgaggggtggtcttctacag-tggggcctctc 479
 |||||
 Db 496 TCCCCCA-CCCTGTCTCTCGATCTATGGGGGTGGCTTACAGTTGGGGCTCTCTC 554
 QY 480 ctgagcgtgacagatggccgctctctgtgtacagccgagagactgtgtgtccat 539
 |||||
 Db 555 CTGGACGTGTACGATGGCCGCTCTTGTGTACAGCCGAGAGACTGTCTGTGTCCAT 614
 QY 540 gaactaccgggtggagcctttggcttctgctgcctgcccgtggggagccgagagcccggg 599
 |||||
 Db 615 GAACTACCGGGTGGAGCCTTTTGGCTTCTTGGCTGCGCGGGAGCCGAGAGG-CCCGGG 673
 QY 600 caatgtgggtctctcg 615
 |||||
 Db 674 CAATGTGGGTCTCCTG 689

RESULT 15

BE535653
 LOCUS 601062390F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3448679 5',
 DEFINITION mRNA sequence..
 ACCESSION BE535653
 VERSION BE535653.1 GI:9764298
 KEYWORDS EST..
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 722)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM8424 row: d column: 24
 High quality sequence stop: 684.
 Location/Qualifiers
 1..722
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3448679"
 /clone_lib="NIH_MGC_10"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site.1: NotI;
 Site.2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.5 kb. Library prepared by Life
 Technologies."

FEATURES

source

1..722
 119 a 198 c 266 g 139 t

Query Match

31.7%; Score 547; DB 10; Length 722;

Best Local Similarity 92.9%; Pred. No. 1..le-94;
 Matches 672; Conservative 0; Mismatches 40; Indels 11; Gaps 9;

QY 478 tcttggagctgtacgatggccgctcttctgttacagccgagagactgtcgtggtgtcc 537
 |||||
 Db 1 TCTTGGAGCTGTACGATGGCCGCTTCTTGGTGTACAGCCGAGAGGACTGTGCTGTGTC 60
 QY 538 atgaactaccgggtggagcctttggcttctcctggccctgccgggggagcagagagcccg 597
 |||||
 Db 61 ATGAACCTACCGGTGGAGCCTTTGGCTTCTTGGCCCTGCCGGGAGCCGAGAGGCCCG 120
 QY 598 gcaatgtgggtctctcctgagatcagagcctggccctcagtggtggtgagagaacgt-ggc 556
 |||||
 Db 121 GGCAATGTGGGTCTCTCTGGATCAGAGCCTGGCCCTCAGTGGGTGGTGCAGGAACTGTGG 180
 QY 657 agccttcgggggtgaccgcacatcagtcagctgtgttgggagagcgggagcgcgcctc 716
 |||||
 Db 181 AGCCTTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCCGGAGCCG-CTC 239
 QY 717 ggtggcatgcacctgctgtccccgccagccggggcctgttccacagggccgtgctgca 776
 |||||
 Db 240 GGTGGGCATGCACCTGCTGTCCCGCCAGCCGGGCCCTGTTCCACAGGGCCGTGCTGCA 299
 QY 777 gagcgtgcccccaatgagacctgggcccacggtgggcatgggagagagcccgctcgagggc 836
 |||||
 Db 300 GAGCGGTGCCCCCAATGGACCTGGGCCACGCTGGGCATGGGAGAGGCCCGCTCGCAGGGC 359
 QY 837 cagcga-gctggccacacctgtgggctgtcctccagggcgccac-tggtgggaatgacaca 894
 |||||
 Db 360 CACGCATGGTGGCCCACTTGTGGGCTGTCTCCAGCGGCACCTTGGTGGGAATGACACA 419
 QY 895 gagctggtgagcctgcttcggacaagcagcagcagcagcagcagcagcagcagcagcagc 954
 |||||
 Db 420 GAGCTGGTAGCCTGTCTCGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 479
 QY 955 gtgctgctcaagaaagcgttctccgggttctccttctccttctccttctccttctcctt 1014
 |||||
 Db 480 GTGCTG-CTCAAGAAAGCGCTCTTCCGGTTCCTCTCGTCCCTGTGGTAGATGGAGACTTC 538
 QY 1015 ctacgtgacacccagagagccctcatcaacgcgagagacttccacggcctgcaggtgctg 1074
 |||||
 Db 539 CTCAGTGACATCCAGAGG-CCTCATCAACGCGGGGAGA--TTCACGGCTGCAGAGTGCTG 595
 QY 1075 gtgggtggtggaagatgagggctcgtatttcttggtttacggggcccccagccttcagc 1134
 |||||
 Db 596 GTGGGTGTTGGTGAAGGATTGAGGGCTCGTATACTGTTAAAGGGGGCCCGCAGCTTCAG- 654
 QY 1135 aaagacaacagctctctcctcagccggccgagttcctcagccgggtgcgggtcggggtt 1194
 |||||
 Db 655 AAAGACACGAAGTCTCTCATCAGC--GGGCGAGTACTTGGCGGGGTGCGGGTTCGGGGTT 712
 QY 1195 ccc 1197
 |||||
 Db 713 CCC 715

Search completed: August 31, 2002, 15:26:22
 Job time: 6485 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:34:03 ; Search time 153.1 Seconds
(without alignments)
2767.588 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atagagcccccgcagtgctc.....aattgtcagcgccacctga 1725

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/1/lna/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/lna/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/lna/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/lna/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1722.4	99.8	3096	2	US-08-318-826A-6
2	1722.4	99.8	3096	2	US-08-370-156-3
3	1722.4	99.8	3096	3	US-08-814-095-3
4	1722	99.8	1845	1	US-07-732-962A-1
5	1722	99.8	1845	5	PCT-US92-06106-1
6	1722	99.8	2256	2	US-08-318-826A-5
7	1722	99.8	2256	2	US-08-370-156-1
8	1722	99.8	2256	3	US-08-814-095-1
9	1722	99.8	3016	2	US-08-318-826A-7
10	1722	99.8	3016	2	US-08-370-156-5
11	1722	99.8	3016	3	US-08-814-095-5
12	1198	69.4	35060	3	US-08-814-095-7
13	380	22.0	2381	2	US-08-318-826A-9
14	380	22.0	2416	2	US-08-318-826A-8
15	380	22.0	2416	4	US-09-334-489-1
16	380	22.0	2416	4	US-09-334-489-2
17	378.4	21.9	2400	6	US15909-13
18	322.8	18.7	2445	6	US15909-9
19	185.6	10.8	764	6	US15909-7
20	142.6	8.3	3018	1	US-08-347-718B-3
21	142.6	8.3	3018	1	US-08-482-262-3
22	141	8.2	3018	6	5200183-1
23	139.4	8.1	2184	1	US-08-445-050-8
24	139.4	8.1	2184	1	US-08-204-691-8
25	139.4	8.1	2428	1	US-08-445-050-1
26	139.4	8.1	2428	1	US-08-204-691-1
27	139.4	8.1	2487	3	US-08-370-223-12

28	134.6	7.8	1907	1	US-08-462-884A-2	Sequence 2, Appl
29	134.6	7.8	1908	1	US-08-461-881B-2	Sequence 2, Appl
30	134.6	7.8	1908	2	US-09-123-960-2	Sequence 2, Appl
31	104.2	6.0	1753	4	US-09-058-260-21	Sequence 21, Appl
32	104.2	6.0	1756	4	US-09-058-260-31	Sequence 31, Appl
33	104.2	6.0	1776	4	US-09-058-260-23	Sequence 23, Appl
34	104.2	6.0	1896	4	US-09-058-260-3	Sequence 3, Appl
35	104.2	6.0	1952	4	US-09-058-260-5	Sequence 5, Appl
36	104.2	6.0	1957	4	US-09-058-260-19	Sequence 19, Appl
37	104.2	6.0	4090	2	US-08-781-802-5	Sequence 5, Appl
38	104.2	6.0	4090	4	US-08-694-078-5	Sequence 5, Appl
39	104.2	6.0	6263	2	US-08-781-802-3	Sequence 3, Appl
40	104.2	6.0	6263	4	US-08-694-078-3	Sequence 3, Appl
41	102.6	5.9	1899	4	US-09-058-260-13	Sequence 13, Appl
42	100.2	5.8	1800	1	US-08-484-815-11	Sequence 11, Appl
43	100.2	5.8	1800	3	US-08-888-949-11	Sequence 11, Appl
44	100.2	5.8	1800	4	US-08-888-950-11	Sequence 11, Appl
45	100.2	5.8	1800	4	US-09-262-758-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-08-318-826A-6

; Sequence 6, Application US/08318826A

; Patent No. 5891725

; GENERAL INFORMATION:

; APPLICANT: Soreq, Hermona

; APPLICANT: Zakut, Haim

; APPLICANT: Eckstein, Fritz

; TITLE OF INVENTION: Synthetic Antisense

; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions

; TITLE OF INVENTION: Containing Them

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kohn & Associates

; STREET: 30500 No. 5891725thwestern Hwy., Suite 410

; CITY: Farmington Hills

; STATE: Michigan

; COUNTRY: US

; ZIP: 48334

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/318,826A

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Kohn, Kenneth I.

; REGISTRATION NUMBER: 30,955

; REFERENCE/DOCKET NUMBER: 2391.00001

; TELEPHONE: (248) 539-5050

; TELEFAX: (248) 539-5055

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3096 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 160...1959

; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370.156
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-307 (Mulford)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3096 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 160..1959
US-08-370-156-3

Query Match 99.8%; Score 1722.4; DB 2; Length 3096;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgagggccccagtgctgctgcacacgcttccctgcttccccactctctccctc 60
DB 160 ATGAGGCCCCCGAGTGCTGCTGCACACGCCCTTCCTGGCTTCCCACTCTCTCCCTC 219
QY 61 ctctcttgctcctgggtgagaggtggtggtgagggccgggaggtacagagctgtg 120
DB 220 CTCTCTGTGCTCTGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGACAGCTGTCTG 279
QY 121 gtgacggtgctggtgggcccgtgcgggcatctgctgaagacccccggggcccctg 180
DB 280 GTGACGGTGCCTGGGGCCGGCTGGGGGCATTCGCTGAAGACCCCGGGGGCCCTGTC 339
QY 181 tctgtcttctgggaccccccttgcggagccaccatgggacccctgccttcttgcca 240
DB 340 TCTGCTTCTGGGATCCCTTTTGGGAGACCCACCATGGGACCCGCTGCTTCTGCCA 399
QY 241 cgggagcccaagcagccttggtaggggtgtagcgtacacattccagagtgctgc 300
DB 400 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGCGCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtgacacacctataccccaggttttgaggccacagagatgtggaaccccaac 360
DB 460 TACCAATATGTGGACACCTATACCAGGTTTGGAGGACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcagagactgctgtacctcaacgtgtgacaccataccccggccctaca 420
DB 520 CGTGAGCTGAGGAGGAGCTGCTGTACCTACAGTGTGACACCATAACCCCGGCCCTACA 579
QY 421 tccccacccccctgctcgtggtatctatggtgggtggttctacagtggtggcctctcc 480
DB 580 TCCCCACACCCCTGCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC 639
QY 481 ttggacgtgtacagtgccgcttcttggttacaggccagagagactgtgtgtgtccatg 540
DB 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCGGAGAGACTGTGTGTGTGTCCATG 699
QY 541 aactaccgggtggagccttggcttcttgccttgccttgcgggagccagagacccccgggc 600
DB 700 ACTACCGGGTGGAGCCCTTGGCTTCTTCCCTGCCCCGGGAGCCGAGAGGCCCGGGC 759
QY 601 aatgtgggtctctggtacagaggtggccctgcagtgagtgagagagacgtggcagcc 660
DB 760 AATGTGGGTCTCTTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACCTGGGAGCC 819

QY 661 ttccgggggtgacccgacatcagtgacgctgttttggggagagcgcgggagccgctcggtg 720
DB 820 TTCCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGAGCCGCTCGGTG 879
QY 721 ggcattcacctgctgtccccgccagcgggcttccacagggcgtgctgcagagc 780
DB 880 GGCATGCACCTGCTGTCCCGCCGACCGGGGCTGTTCACAGGGCCGCTGCTGCAGAGC 939
QY 781 ggtgcccccaatggacccttggccacggtgggcatgggagagcccgctgcagggccacg 840
DB 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGCATGGGAGAGGCCCGCTGCAGGGCCACG 999
QY 841 cagctggcccaacctgtgggctgtctccagcggcactggtgggaatgacacagagctg 900
DB 1000 CAGCTGGCCACCTTGTGGGCTGTCTCCAGCGGCACATGGTGGGAATGACACAGAGCTG 1059
QY 901 gtgacctgcttcggacacgacagcagcgaggtcctcgtgtaaccacgaaatggcagctg 960
DB 1060 GTAGCCTGCTTCGGACACGACGACGAGGTCTCTGCTGNACCCACGATGGCACGTGCTG 1119
QY 961 cctcaagaaagcgttctccgggttctcctctgctgctgtgtagatggagacttctcctcag 1020
DB 1120 CCTCAAGAAAGCGCTCTCCGGTCTCTCTCGTGCCTGTGTAGATGGAGACTTCTCTCAGT 1179
QY 1021 gacaccccaagagccctcctcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
DB 1180 GACACCCGAGAGCCCTCTCAACACGCGGAGACTTCCACGGCCTGCAGGTGCTGTGGGT 1239
QY 1081 gtggtgaagatgagggctgctatttctggtttacggggcccgagcttcagcaagac 1140
DB 1240 GTGGTGAAGATGAGGGCTGCTATTTTCTGTTTACGGGGCCCGAGCTTTCACGAAGAC 1399
QY 1141 aacgagctctctcatcagccgggcccaggttctctgcccgggtgctgggttcggttccccag 1200
DB 1300 AACGAGCTCTCTCATACGCCGGCCGAGTTCTTGCGCGGGGTGCGGGTTCGCCAG 1359
QY 1201 gtaagtgcacctggcagccagagctgtggtcctgcataacacagactggtcgtccag 1260
DB 1360 GTAAAGTGACCTGGCAGCCGAGGCTGTGCTCTGCATTTACACAGACTGGCTGCATCCGAG 1419
QY 1261 gaccggcagcctgagggagggcctgagcgtggtggtgggagccacacatgctgctgc 1320
DB 1420 GACCCGGCAGCCTGAGGAGGGCCCTGAGCGATGTGTGGGGCCACCAATGTGTGTGTC 1479
QY 1321 cccgtggccacagctggtggtggcagctggtgcccaggggtgcccgggtctctacgctacgtc 1380
DB 1480 CCGTGGCCACAGCTGGCTGGCGACCTGGCTGCCAGGGTGCCCGGGTCTACGCTACGTC 1539
QY 1381 ttgaaacacgctgcttccacgctcctcgtggccctgtggtgggggtgccccagcgtac 1440
DB 1540 TTGTGAACACCGTGTCTTCCACGCTCTCTGGCCCTCTGGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcagttcatctttgggataccccctggacccctctgaaactacacagcgagagag 1500
DB 1600 GAGATCGAGTTTCATCTTTGGGATCCCTTGACCCCTCTCTGAAACTACACGGGAGAGGAG 1559
QY 1501 aaaaattctgccacgactgatgcatactggggccaaactttgccgcacagggggtcccc 1560
DB 1660 AAAATCTTGCCCCAGCAGCTGATGCGATACTGGGCCAACTTTGCCGACACAGGGATCCC 1719
QY 1561 aatgacccccagagacccccaaagccccacaaatggcccccgtaacacgcgggggtctacag 1620
DB 1720 AATGACCCCGAGACCCCAAGGCCCCCAATGGCCCGCTACACGGCGGGGTCTACGAG 1779
QY 1621 tacgttagctgagcctgcggcgtgaggtgctgggggggctgcgcgccagggcctgc 1680
DB 1780 TAGCTTAGTCTGACCTGCGGCCGCTGGAGGTGCGGGGGGTGCGGCCCGAGGGCTGC 1839
QY 1681 gcctctggaacgcgttctctccccaaatgctcagcgccacctg 1724
DB 1840 GCCTTCTGGAACCGCTTCTCTCCCAATTTGCTCAGCGCCACCG 1883

RESULT 3
US-08-814-095-3
Sequence 3, Application US/08814095
Patent No. 6025183
GENERAL INFORMATION:
APPLICANT: Soreq, Hermona
APPLICANT: Zukut, Haim
APPLICANT: Shani, Moshe
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 No. 6025183thwestern Highway, Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3096 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Alternatively spliced Ache
DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
DESCRIPTION: of intron 4 (readthrough)."
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 160..1959
US-08-814-095-3

Query Match 99.8%; Score 1722.4; DB 3; Length 3096;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctgtgcacacgccttccttggtctcccaactctctctc 60
DB 160 ATAGGCCCCCGAGTGTCTGTCTGCACACGCTTCCCTGGCTTCCCACTCTCTC 219
QY 61 ctctcttggtctctggtgagagtggtgggtgagggccgggagagatgcagagctgtg 120
DB 220 CTCTCTGTGGTCTGTGGTGTGAGAGTGGGGCTGTAGGGCCGGGAGGATGCAGAGCTGTG 279
QY 121 gtacaggtgcgtggggccggctgcggggcattcgcctgaagacccccggggccctgtc 180
DB 280 GTACGGTGTGGGGCCGGCTGTGGGGGCAATTCGCTTGAAGACCCCGGGGGCCCTGTC 339
QY 181 tctgtcttcttggtaccccttttgcggagccacccatggagacccctgtcttcttcca 240
DB 340 TCTGCTTCTTGGGCACTCCCTTTTGGGAGGCCACCCATGGGAGCCCGCTTCTGCCA 399
QY 241 ccggagcccaagcagccttggttcaggggtgtagacgtacacaccccttcagagtgctgc 300

DB 400 CCGAGGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGTACAACTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctataccaggttttgaggccaccgagatgtggaaccccaac 360
DB 460 TACCAATATGTGGACACCCATATACCCAGGTTTTCAGGGCCACCGAGATGTGAACCCCAAC 519
QY 361 cgtgagctgagcaggactgctgtacctaacaactgtggacacacataccccccggcctaca 420
DB 520 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCTTACA 579
QY 421 tccccacccctgtcctcgtctggtatctatggtggtggtctctacagtggggcctctcc 480
DB 580 TCCCCACCCCTGTCTCTGTGGATCTATGGGGGTGCTTCTACAGTGGGGCTTCTCTCC 639
QY 481 ttgacgtgtacgatggccgcttcttggtacagcccgagagactgtcgtggttccatg 540
DB 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGCCGAGGAGACTGTCTGTGTCTCATG 699
QY 541 aactaccgggtgggagcctttgcttccctggccctgcggggagccgagagcccccgggc 600
DB 700 AACTACCGGGTGGGAGCCTTTGGCTTCTGCGCCCTGCGGGGAGCCGAGAGGCCCGCGGC 759
QY 601 aatgtgggtcctcgtgatacagagctggccctgcagtggtggtgaggaacactgtggagcc 660
DB 760 AATGTGGGTCTCTGTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAACTGGCCAGCC 819
QY 661 ttccgggggtgacccgacatcagtgacgtctgttggggagagcgcggagccgcctcgtg 720
DB 820 TTCCGGGGGTGACCCGACATCAGTACGCTGTCTTGGGAGAGCCGCGGAGCCCTCTCGGTG 879
QY 721 ggcatgcacctgtgtcccccggccagcggggcctgttccacagggcgctgtcagagc 780
DB 880 GGCATGCACCTGCTGTCCCGCCCGCCAGCGGGGCTGTTCACAGGGCGCTGCTGCAGAGC 939
QY 781 ggtgcccccaaatggacccctgggcccacggtgggcatgggagagggccgtcgagggccacg 840
DB 940 GGTGCCCCCAATGGACCTTGGGCCACGCTGGGCATGGGAGAGGCCCTTCGAGGGCCACG 999
QY 841 cagctggccaccctgtgggctgtctccagggcgccactggtgggaatgcacacagactg 900
DB 1000 CAGCTGCCCCACCTTGTGGCTGTCTCCAGCGCGACTGTGTGGGAATGACACAGAGCTG 1059
QY 901 gtacgtcctcgttcgacacacacgcagcagctcctgtgtaaccacgaatgcagcgtgctg 960
DB 1060 GTAGCCTGCTTTCGGACACGACGACGAGCTCTGTGTGTAACCAAGAAATGCGAGTGTCTG 1119
QY 961 cctcaagaaagcgtcttccggttctcctcgtcgtgtgtgtagatgagagcttctcagt 1020
DB 1120 CCTCAAGAAAGCGTCTTCCGGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
QY 1021 gacacccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
DB 1180 GACACCCAGAGGGCCCTCATCAACGCGGAGACTTCCACGGCTGTGAGGTGTGTGTGTGTGT 1239
QY 1081 gtggtgaagatgagggctcgtatatttctggtttacggggcccccaggtcttcagcaagac 1140
DB 1240 GTGTTGAAGATGAGGGCTGCTATTTCTGTTTACGGGGCCCCCAGGCTTTCAGCAAAAGAC 1299
QY 1141 aacgagctctcatcagccgggcccaggttctcgtgcccgggtgcgggttcggggttccccag 1200
DB 1300 AACGAGTCTCTCATCAACGCGGGCGGAGTCTCTGCGGGGTGCGGGGTTCGGGGTTCGCCAG 1359
QY 1201 gtaagtgcacctggcagccgggctgtgttctcgtcatlacacagactggctgcatccccag 1260
DB 1360 GTAAGTCACTGGCAGCCGAGGCTGTGTCTGCTCATTACACAGACTGGCTGCATCCCGAG 1419
QY 1261 gacccggcagcctcgtgagggggccctgagcgtatgtgtggtgggagaccacaaatgtcgtgctgc 1320
DB 1420 GACCCGACGCTGTGAGGGAGGCCCTGTGAGGATGTGTGGGGAGCCACCAATGTCTGTGTGC 1479
QY 1321 ccggtggcccgagcgtgctggcgactggtgcccagaggtgcccgggtctcagcgcactgctc 1380


```
Db 1021 GACACCCAGAGGCCCTCATACACGGGAGACTTCACAGGCCCTGCAAGGTGCTGGTGGT 1080
Qy 1081 gtggtgaagatgaggtctgtatttctgtttacaggggccccaggtcttcagcaaaagac 1140
Db 1081 GTGGTGAAGGATGAGGCTCGTATTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1140
Qy 1141 aacgagtctctacagcgggcccagagttctctgcccgggtgcgggttcgggttccccag 1200
Db 1141 AACGAGTCTCTCATCAGCGGGCGGAGTCTCTGGCCGGGTGCGGGTTCGCCAG 1200
Qy 1201 gtaagtacacctgagcggaggtgtgtctgtgattacacagactggtctgacccccag 1260
Db 1201 GTAAGTACCTTGGCAGCGGAGGCTGTGGTCTGCAATTTACACAGACTGGCTGCATCCCGAG 1260
Qy 1261 gacccggcagcctgagggagccctgagcagatgtgtggcgaccacaatgtcgtgtgc 1320
Db 1261 GACCCGGCAGCCCTGAGGAGGCCCTGAGCGATGTGTGGCGGACCAATGTCTGTGC 1320
Qy 1321 cccgtggccagctgggtgggagactgggtgcccaggggtgcccgtctacgcctacgtc 1380
Db 1321 CCCGTGGCCAGCTGGCTGGCGGACTGGCTGCCAGGGTGGCCGGGTCTACGCCCTACGTC 1380
Qy 1381 ttgaaacacgtgttccacagctctctgcccctgtggatgggtgccccacaggtac 1440
Db 1381 TTTGAAACACCGTGTTCACAGCTCTCTGGCCCTGTGGATGGGGGTGCCCCACGGGTAC 1440
Qy 1441 gagatcagttcttggatccccctggaacccctctcgaatactacacggcagagag 1500
Db 1441 GAGATCAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGAG 1500
Qy 1501 aaatcttgcggcagctgactgactgatactggtggccaaacttggcccacaggggtatccc 1560
Db 1501 AAAATCTTGGCCAGCAGCTGATGCGATCTGCGCCAACTTTGCCCGCACAGGGGATCCC 1560
Qy 1561 aatgagccccagacccaaagcccccaaatgcccccgatgcccccgatgcccccgag 1620
Db 1561 AATGAGCCCCGAGACCCAAAGCCCCCAATGCCCCCGCTACACGGCGGGGGCTCAGCAG 1620
Qy 1621 tacgttagctggacctgcggcgtgaggtgaggtgaggtgaggtgaggtgaggtgaggtg 1680
Db 1621 TACGTTAGCTGGACCTGGCGCGCTGAGGTGGCGGGGGCTGCGCGCCAGGCGCTGC 1680
Qy 1681 gcttctggaacctctctctctctctctctctctctctctctctctctctctctctct 1720
Db 1681 GCCTTCTGGAACCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
```

RESULT 5

PCT-US92-06106-1

; Sequence 1, Application PC/TUS9206106

; GENERAL INFORMATION:

; APPLICANT: Fischer, Melir

; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN

; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White, Esq.

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/06106

; FILING DATE: 19920722

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

```
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
; PCT-US92-06106-1
```

Query Match 99.8%; Score 1722; DB 5; Length 1845;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 atgaggccccgcagtgctgtctgcacacgaccttccctggttccccactcttctc 60
Db 1 ATGAGGCCCCGCGAGTGTCTGTGCACACGCGCTTCCCTGGCTTCCCGACTCTTCTCTC 60
Qy 61 ctctcttggctcttggatggagagtgagggtgagggcgagggcgagggatgagagctgtg 120
Db 61 CTCTCTTGGCTTCTTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGATGCGAGAGCTGTCTG 120
Qy 121 gtacggtgtggtggggcggtcggtcggttcgctgaagacccccggggggccctgtc 180
Db 121 GTGACGCTGCTGGGGCGCGGCTCGCGGGCATTCGCTGAAGACCCCGGGGGCCCTGTCTG 180
Qy 181 tctgttcttgggcacatcccttggggagccacccatgggagccccctgctttcttccca 240
Db 181 TCTGCTTTCCTGGGATCCCTTTTGGGAGACACCCATGGGAGCCCTCGCTTCTTCTGCA 240
Qy 241 ccgagccccaaagccttggatgaggtgtgtagacgtacacgttccagagtgctgc 300
Db 241 CCGAGGCCAAGCAGCCTTGGTCAAGGGGTGTACGCTACACCTTCCAGAGTGTCTGCG 300
Qy 301 taccaatagtgtggacacccctataccaggttttggggcaccgagatgtggaaccccaac 360
Db 301 TACCAATATGTGGACACCTATACCCAGGTTTGGGGCACCCGAGATGTGGAACCCCAAC 360
Qy 361 cgtgagctgagcaggagctgctgtacacacgtgtggacacacaccccccgccctaca 420
Db 361 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCTTACA 420
Qy 421 tccccccccctgtcctgtctgtatctatggggtggtctctcagtggggctctctcc 480
Db 421 TCCCCCACCCCTGCTCTGCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCCCTCTCTCC 480
Qy 481 ttggacgtgtacagtgccgcttcttggacagcccgagagagactgtgtgtgtccatg 540
Db 481 TTGGACGTGTACGATGCCGCTTCTTGTACAGCCCGCAGAGAGACTGTGTGTGTCCATG 540
Qy 541 aactaccgggtggagccttggcttctcctgcccccgaggggagccgagagggccccgggc 600
Db 541 AACTACCGGGTGGAGGCTTTGGCTTCTTGGCTTCTGGGGAGCCGAGAGAGCCCGGGGC 600
Qy 601 aatgtggtctcctggtatcagagcgtgcccctcagtggtgtgagagagcgtgagcagcc 660
Db 601 AATGTGGTCTCTGATCAGAGGCTGGCCCTGAGTGGGTGGGTGAGAGAACTGTGGCAGCC 660
Qy 661 ttccgggggtaccgcacatcagtgacgtgttttggggagagcgcggagccgctcgtgtg 720
Db 661 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCGCGCTCGGTG 720
Qy 721 ggcattgacattgtgtcccccgccccagccggggcgtgttccccagggcggtgtgtcagagc 780
```

```
Db 721 GGCATGACCTGCTCTCCCGCCGAGCGGGCCCTGTTCCACAGCGCGTGTGCAGACC 780
Qy 781 ggtgcccccaatggacctggccacggtgggcatgggagagcccgctgcagggccacg 840
Db 781 ggtgcccccaatggacctggccacggtgggcatgggagagcccgctgcagggccacg 840
Qy 841 cagctggccccacctgtggctgtctccaggcgccactggtgggaatgacacagagctg 900
Db 841 CAGCTGGGCCCCACTTGTGGCTGTCTCTCCAGCGGCACATGGTGGGAATGACACAGCTG 900
Qy 901 gtgactgccttcggacacgaccagcgaggtccctgggtgaaccacgaaaggacgtgctg 960
Db 901 GTAGCCTGCCCTTCGGACACGACGCGCAGGTCCTTGGTGAACACCAAGATGGCACGTG 960
Qy 961 cctcaagaaagcgtcttcocggttctcctctgtgctctgtgtgtagatggagacttctcagt 1020
Db 961 CCTCAAGAAAGCGTCTCCGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 gacacccagagggcctcctacacacgggagagacttccacggcctcaggtgctggtggt 1080
Db 1021 GACACCCAGAGGGCCTCATCAACCGGGAGACTTCCACGGGCTCGAGGTGCTGGTGGGT 1080
Qy 1081 gtgggtgaaggatgaggtctgtatcttctgtttacggggcccccaggtctcagcaagac 1140
Db 1081 GTGGGTGAAGGATGAGGCTGCTGATTTCTGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1140
Qy 1141 aacgagttctctacagcgggcccagttcctggcggggtgctggggttcccccag 1200
Db 1141 AACGAGTCTCTCATCAGCGGGCCGAGTTCCTGGCGGGGTGCGGGTTCCTCCAG 1200
Qy 1201 gtaagtgcctggcagcagaggtgctgtctctgctgattacacagactgctgctcagcag 1260
Db 1201 GTAAGTGACCTGGCAGCGCAGGCTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Qy 1261 gaccggcacgctgagagagccctgagcagatgtgggtggcgacacaaatcgtgtgc 1320
Db 1261 GACCGGGCAGCGCTGAGGAGGCCCTGAGGATGTGGTGGGCGACCAATGTCTGTGTC 1320
Qy 1321 cccgtggccccagctggctggggaactgctgcccgggtgcccgggttacccctacgtc 1380
Db 1321 CCGGTGGCCACGCTGGCTGGGCGACTGGCTGCCAGGGTGCCTGGCTTACGCCCTACGTC 1380
Qy 1381 tttagaacctgtctcagcgtctccctggccctgcccctgtgagtggtggccacagcctac 1440
Db 1381 TTGTGAACACCGTGTCTTCACGCTCTCTGTGGCCCCCTGTGGATGGGGGTGCCCCACG 1440
Qy 1441 gagatcagattcattcttggggtacccctggaccctctcgaactacacagcagagag 1500
Db 1441 GAGATCGAGTTCATCTTTGGGATCCCTCTGGACCCCTCTCGAAACTACACGGCAGAGG 1500
Qy 1501 aaaaattcttcgcccagcagactgtagatgctggtggcgaactttgcccgcacagggatccc 1560
Db 1501 AAAAATCTTGCCCGCAGGACTGATGATGACTGGGCAACTTTTGGCCGCACAGGGATCC 1560
Qy 1561 aatgagccccagagaccccccaagcccccaaatggcccccgatacagcgggggctcagcag 1620
Db 1561 AATGAGCCCCAGAGACCCCAAGCCCCCAATGGCCCCCGGTACACGGCGGGGCTCAGCAG 1620
Qy 1621 tacgttagtctgagctgcggccgctggaggtgaggggtgagggggtgctgcgccccagcctgc 1680
Db 1621 TACGTTAGTCTGAGACTCTGGGCGCTGTGAGGTGCGGGGGGCTGCGGCCCCAGGCGCTGC 1680
Qy 1681 gctcttggaacgcttctcctcccaaatgtcagcgccacc 1722
Db 1681 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1722
```

RESULT 6

US-08-318-826A-5
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:

APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Eckstein, Fritz
TITLE OF INVENTION: Synthetic Antisense
TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
TITLE OF INVENTION: Containing Them
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 5891725thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,826A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 2391.00001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2256 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
OTHER INFORMATION: 3, 4 and 6"

Query Match 99.8%; Score 1722; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 atgagggccccccagtgctgtgcacagccttccctggcttccccactcccttctcctc 60
Db 160 ATGAGGCCCCCAGTGTCTGTGCACACGCTTCCCTGGCTTCCCACTCTCTCTCTC 219
Qy 61 ctctctgtgctcctgggtggagagtggtgggtgagggccggagagatgagagctgctg 120
Db 220 CTCTCTGTGCTCTCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTG 279
Qy 121 gtgacggttcgtggggccggcgtggtggcgttcctgtaagacccccggggccctgtc 180
Db 280 GTGACGGTTCGTGGGGCCGGCTGCGGGGCATTCGGCTTGAAGACCCCGGGGCCCTG 339
Qy 181 tctgtcttctcctgggcatcccttgcggagaccacccatggaccctgctcttctgcca 240
Db 340 TCTGCTTCTCTGGGCATCCCTTTGGGGAGCCACCCTATGGGACCCCGCTTCTTCCCA 399
Qy 241 ccggagcccccaagcagccttgggtcaggggtggttagacgtacacacctccagagtgctgc 300
Db 400 CCGAGCCCCAAGCAGCGCTTGGTCAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 459
Qy 301 taccaatatgtggacacccctataccagatgttggagggccacagagatgtggaaacccaac 360
Db 301 taccaatatgtggacacccctataccagatgttggagggccacagagatgtggaaacccaac 360
```

Db 460 TACCATATGTGGACACCCCTATACCCAGGTTTGTAGGGGACCCAGAGATGTGGAAACCCCAAC 519
QY 361 cgtgagctgagcagagactcctctacactcaacgctgagacacatacccccgcctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCGTGACCTCAACGCTGTGGACACCATACCCCGGCTTACA 579
QY 421 tccccaccctctcctcgtctcgtgatactatg9999t99ctctacagtg9999ctctcc 480
Db 580 .TCCCCACCCCTGCTCTGCTGTGATCTATGGGGTGGCTTCTACAGTGGGGGCTCCTCC 639
QY 481 ttgagctgtacatagccctctcttggtacagccgagagagactgtctggtgtccatg 540
Db 640 TTGAGCTGTACGATGCCCCCTTCTTGGTACAGCCGAGAGACTGTGCTGGTGTCCATG 699
QY 541 aactaccgggtg99agcctttt99cttctcctgcccctgcccgggagccgagagcccg99c 600
Db 700 AACTACCGGTGGAGGCTTTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
QY 601 aatgt9999tctcctgagatagag9999t9999t9999t9999t9999t9999t9999t 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGGAGGAAAGCTGGCAGCC 819
QY 661 ttcg9999t9999t9999t9999t9999t9999t9999t9999t9999t9999t9999t 720
Db 820 TTGCGGGGTACCCGACATCAGTGACCTGTGTTGGGAGAGCGCGGAGCGGCTCGGTG 879
QY 721 ggcacgacactgtgtctcccccagccggg9999t9999t9999t9999t9999t9999t 780
Db 880 GGCATGACCTGTGTCTCCCGCCAGCGGGGCTGTGTTCCACAGGGGCGTGTGCGAGC 939
QY 781 ggt99 840
Db 940 GGTGCCCCCAATGACCTGGCCAGCGTGGGATGGAGAGGCCCGCTGCGAGGGCCAGC 999
QY 841 cagctggccacactgtgggctgtcctccagcgagcactgtggaatgacacagactg 900
Db 1000 CAGCTGCCCCACCTGTGGGCTGTCTCCAGCGGCACTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtgacgtgctctggacacagcagcgaggtcctggtgaaccacgaatggcagctgtg 960
Db 1060 GTAGCTGCTCTCGGACACACAGCAGCGAGGTCTGTGTGAACACAGGAATGGCAGTGTG 1119
QY 961 cctca999 1020
Db 1120 CTTCAAGAAAGCTGTCTCCGGTGTCTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
QY 1021 gacacccagagccctcatcaacg99999999999999999999999999999999999999 1080
Db 1180 GACACCCAGAGGCCCTCATCAAGCGGGAGACTTCCACGGGCTGCGAGGTGTGTGTGTGTGT 1239
QY 1081 gtggtgaag99 1140
Db 1240 GTGTGAAGGATGAGGCTCGTATTTCTGTGTTACGGGCCCCAGGCTTTCAGCAAAAGC 1299
QY 1141 aacgagctctcatcaacg999 1200
Db 1300 AACGAGTCTCTATCAACCGGGGCGAGTTCCTGCGGGGTGCGGGGTGCGGGGTTCGCCAG 1359
QY 1201 gtaagtacactgacgagcgtgtggtcctgcatcacagactggtcgtatcccgag 1260
Db 1360 GTAAGTGACCTGGACCGGAGGCTGTGGTCTGATTTACACACTGCTGTGATTCGCGAG 1419
QY 1261 gacccg999 1320
Db 1420 GACCCGGCAGCCTGAGGGAGGCCCCGTGAGCGATGTGTGTGGCGACACCAATGTCTGTGTC 1479
QY 1321 ccg999 1380
Db 1480 CCGGTGGCCAGCTGGGTGGGCGACTGGCTGCCAGGGTCCCGGGTCTACGCCCTACGTC 1539
QY 1381 ttg999 1440
Db 1540 TTTGAACACCGCTGTTCCACGCTCTCTCTGCCCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1599

QY 1441 gagatcaggttcattctttgggataccccctggacccccctctcgaactacacg9999999 1500
Db 1600 GAGATCAGTTCATCTTTGGATCCCTCCCTGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaaattc99 1560
Db 1660 AAAATCTTCCCGCAGCAGCTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCC 1719
QY 1561 aatgagcccgagaccccaagggcccaaatggcccccgtacacg99999999999999999999 1620
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagctgtgacactgc99 1680
Db 1780 TAGCTTACTGTGACCTGCGCGCTGAGGTGCGGGGGGCTGCGCGCCAGGCCTGC 1839
QY 1681 gccttctggaacgctctctcccaaaattgctcagc99999999999999999999999999 1722
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1881

RESULT 7

US-08-370-156-1
; Sequence 1, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-156-1

Query Match 99.8%; Score 1722; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgag999 60
Db 160 ATAGGCCCCCGCAGTGTGCTGTGCACACGCGCTTCCCTGGCTTCCCGACATCTCTCTCCTC 219

TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "ACHE gene comprising exons
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-814-095-1

Query Match 99.88; Score 1722; DB 3; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagggcccccagtgctgtgcacacgcttccctggcttcccactcttctctc 60
Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCCACTCTTCTCTC 219
Qy 61 ctctcttggtctcctgggtgaggtggtgggtgaggtgaggtgaggtgaggtg 120
Db 220 CTCTCTGCTCTCTGCTGGTGGAGAGTGGGGCTGAGGGCCGGAGAGTGCAGAGCTGCTG 279
Qy 121 gtacagtggtggggccgctcggggcattcgctgaagacccccggggccctctc 180
Db 280 GTGACGGTGTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCCTGTC 339
Qy 181 tctgtttcttgggcatcccttggagaccacccatgggagcccgctgctttctgcca 240
Db 340 TCTGCTTCTTGGGATCCCTTTGGGAGCCACCCATGGACCCGCTTCGCTTCTGCGCA 399
Qy 241 ccgagaccagcagccttggtcaggggtgtagacgtacacaccccttcaggtgtctgc 300
Db 400 CCGAGGCCAAGCAGCCTTGGTCAAGGGGTGTAGACGCTTACAGCTTCCAGAGTGTCTGC 459
Qy 301 taccatattgtggacacctataccaggttttggagccagagatgtgagccccaac 360
Db 460 TACCATAATGTGGACACCTTATACCAAGTGTGAGGGCACCGAGATGTGAACCCCAAC 519
Qy 361 cgtgagctgagcagagactgctgtacatcaacgtgtgacacacaccccgccctaca 420
Db 520 CGTCAGCTGAGCAGGAGTGTCTTACCTCAACGTGTGGACACCATACCCCGGCTTACA 579
Qy 421 tccccccccctgtcctgtgtgatctatgggggtgggttctacagtggggcctctcc 480
Db 580 TCCCCACCCCTGTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCCTCC 639
Qy 481 ttgacatgtacgatgccccttctgtacagccagagagactgtcgtgttcattg 540
Db 640 TTGACAGTGTACGATGGCCGCTTCTGGTACAGCCCGAGAGAGTGTCTGCTGTCTCATG 699
Qy 541 aactacgggtgggagccttggcttctcctgcccctgcccggagcagagagcccgccg 600
Db 700 AACTACCGGGTGGAGCTTTGGCTTCTTGGCTTCTGCGCCCTGCGGGGAGCCGAGAGCC 759
Qy 601 aatgtgggtctcttgatcagaggtctggcctcagtgagtggtgaggtgagcgtggagcc 660
Db 760 AATGTGGGTCTCTCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATG 819
Qy 661 ttccgggtgagccgacatcagtgacgtgtttgggtgagagcgcgagcgcctcgtg 720
Db 820 TTCCGGGGTGACCCGATCATGATGATGATGATGATGATGATGATGATGATGATGATG 879
Qy 721 ggcattgacattgtctcccgccagcggggcctgttccacagggcgtgctgctcagagc 780
Db 880 GGCATGACCTGCTGTCTCCCGCCAGCGGGGCTGTCTTCCACAGGGCCGTGCTGCAGAGC 939

Qy 781 ggtgcccccaatgacacctgggcccacggtgggcaatgggagagggccctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTGGCCACGCTGGGCATGGAGAGGGCCCTGCAGAGGCCACG 999
Qy 841 cagctggcccaacctgtgggtgtcctccagcggcactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACCTTGTGGGTGTCTCTCCAGCGGCACATGTGGGAATGACACAGAGCTG 1059
Qy 901 gtacgtctccttcggacacgaccagcaggtcctgggtgaaccacgaatgcacgtgctg 960
Db 1060 GTAGCTGCTTGGACACGACGAGGCTGCTGGTGAACACGAATGACAGCTGCTGCTG 1119
Qy 961 cctcaagaagcgtctcttccgggtctcttctcgtgctgtggtagatggggaacttctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTCTCTCTTCTGCTGTGGTGTAGATGGAGACTTCTCTCAGT 1179
Qy 1021 gacccccagaggccctcattcaacgcggggagacttccacggcctgcaggtgctggtgggt 1080
Db 1180 GACACCCCGAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTTCAGAGTGTCTGGGT 1239
Qy 1081 gtggtgaaggtatgaggtcgtctatttctggttttacggggcccgaggtcttcagcaagac 1140
Db 1240 GTGGTGAAGGATGAGGCTCGTATTTCTGTTTACGGGGGCCCGAGGCTTCAGCAGAGAC 1299
Qy 1141 aacgagttctctcatcagcggggccgagttcctgcccgggtgctgggtggtgttccccag 1200
Db 1300 AACGAGTCTCTCATCAGCGGGCGAGTTCCTGGCGGGGTGCGGGTTCGGGTTCGCCAG 1359
Qy 1201 gtaagtgaacctggcagccgaggtgtggtcctgcaatcacagactgggtcgtcaccagag 1260
Db 1360 GTAAAGTACCTTGGCAGCGGAGGCTGTGGTCTCTGCTATACACAGACTTGGCTGCATCCCGAG 1419
Qy 1261 gaccggcagccttgagggagcctgagcgtatgtggggcgaccacaaatgctgtctg 1320
Db 1420 GACCCGCGACGCTTGGAGGAGGCCCTGAGCGATGTGGTGGCGACCACTGCTGCTGTGC 1479
Qy 1321 cccgtggccagctggtggtgagcactggtggtggtggtggtggtggtggtggtggtggt 1380
Db 1480 CCCGTGGCCAGCTTGGCTGGCGGAGTGGCTGCCAGGGGTGCCGGGTCTACGCTACGTC 1539
Qy 1381 tttaaacacgtgtctccagcctcctgcccctgtggtggtggtggtggtggtggtggtggt 1440
Db 1540 TTTGAACACCTGCTCTCCAGCTCTCTCTGGCCCTGTGGATGGGGGTGCCACGGGTAC 1599
Qy 1441 gagatcagttcatcttttgggtacccctggaacccctctcgaaactacacggcagagag 1500
Db 1600 GAGATCAGTTTCACTTTGGGATCCCTTGGACCCCTCTCGAANACTACACGGCAGAGGAG 1659
Qy 1501 aaaaatcttgcggcagcagctgatactggtggtggtggtggtggtggtggtggtggtggt 1560
Db 1660 AAAATCTTGGCCAGCAGCTGATGCGATACCTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
Qy 1561 aatgagccccagaccacaaagcccccaaatggccccgtacacgcggggggtctcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGCCCCCAATGGCCCCCGTACACGGCGGGGGCTTCAGCAG 1779
Qy 1621 tacgttagctgagcctgagcgtggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtg 1680
Db 1780 TACGTTAGCTGAGCTTGGCGGCTGAGGTGCGGGGGGCTGCGGGCCCGCACAGGGCCTGC 1839
Qy 1681 gcttctggaacgccttctctcccccaaatgtctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCCAAATTTGCTCAGCGCCACC 1881

RESULT 9

US-08-318-826A-7
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz

; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,836A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160...2010
; OTHER INFORMATION: /note= "Splice Variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and 6"
; US-08-318-826A-7

Query Match 99.8%; Score 1722; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagggcccccagtgctgtctgacacagcccttccctggcttcccccactccttctctc 60
Db 160 ATGAGGCCCCCAGTGTCTGCTGCACACGCCCTTCCTGGCTTCCCCACCTCTCTC 219

Qy 61 ctctctgtctctgggtgagagagtggggctgagggccggagggatgagagctgctg 120
Db 220 CTCTCTGTGCTCTGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 279

Qy 121 gtgacggtgctggggcgccgctgaggcattcgcctgaagaccccccggggccctgtc 180
Db 280 GTGACGGTGTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCCTGTC 339

Qy 181 tctgttctctgggacatcccccttggcggaccacccatgggaccccgctcttctgcga 240
Db 340 TCTGCTTCTTGGGCATCCCTTTTGGGAGCCACCCATGGGACCCCGCTCGCTTCTGCCA 399

Qy 241 ccggagcccaagacgcttggtcaggggtggtagagcgtacacaccttccagagtgtcgc 300
Db 400 CCGAGCCCCAAGCAGCGCTTGGTTCAGGGGTGGTAGCGCTACAACCTTCCAGAGTGTGC 459

Qy 301 taccaatatgtgacacccctatccccaggtttttaggggaccagagatgttgaaacccaac 360
Db 460 TACCANTATGTGACACCCCTATACCCAGGTTTGTAGGGGCACCGAGATGTGGNAACCCCAAC 519

Qy 361 cgtgagctgagcagagactcctgtacctcaacgctgtggacacatacccccgccctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCGTACCTCAAGCTGTGGACACCATACCCCGGCCCTACA 579

Qy 421 tccccacacccctctcctcgtctgagatatg9999tggcttctacagtg9999cctctcc 480
Db 580 TCCCCACACCCCTGTCTCTGCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTC 639

Qy 481 ttggacggtgtacgatggccgcttcttgggtacagccgagagagagactgtgctggttccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699

Qy 541 aactacgggtg999agccttcttggcttccctggccctgc999ggagccgagagggcccccggc 600
Db 700 AACTACGGGTGGGAGCGCTTGTGGCTTCTGGCCCTGCCGGGAGCCGAGAGGCCCGCCG 759

Qy 601 aatgtgggtctcctggtacagagagcctgagcctgcagtg999tgcaggagagacgtg99cagcc 660
Db 760 AATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAAACGTGGCAGCC 819

Qy 661 ttccgggggtgaccccgacatcagtgacgctgttttggggagagcgcg99gagccgctcggtg 720
Db 820 TTCCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCGCCCTCCGGTG 879

Qy 721 ggcctgacactgtgttccccgccagcc99gggctgtttccacaggg99cgtgctgcagagc 780
Db 880 GGCATGCACCTGTGTCCCGCCAGCGCGGGCCGTGTTCACAGGGCCGCTGCTGCAGAGC 939

Qy 781 ggtgcccccaaatgagaccctggccacggtg999catgggagagagcc99cgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTGGGCCACGCTGGGCATGGGAGAGGCCCTCGCAGGGCCACG 999

Qy 841 cagctggcccaacctgtgggctgtccctccagggcg99cactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCACTGTGGGCTGTCTTCCAGCGGGGCACTGTTGGGAATGACACAGAGCTG 1059

Qy 901 gtgagctgccttcggacacagcagcaggttccctg999tgaacacacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCTTCGGACACAGCAGCGCAGGTCCTGTTGAACACGAAATGCGACGCTGCTG 1119

Qy 961 cctcaagaaagcgtcttcgggttctcctcgtcctcgtgctcgtgtagatgagagacttccagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTCTCTCTGCTGTGTAGATGGAGACTTTCCTCAGT 1179

Qy 1021 gacacccagagccctcatcaacg99ggagacttccacggcctgcaggtgctggtggtg 1080
Db 1180 GACACCCACAGGCCCTCATCAACGGCGGGAGACTTCCACGGCCCTGCAGGCTGTGTTGGT 1239

Qy 1081 gtggtgaagagatgagggctcgtatttctggtttacggggggcccccag99ccttcagcaagac 1140
Db 1240 GTGGTGAAGCATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCGCAGGCTTTCAGCAAGAC 1299

Qy 1141 aacgagctctctcatcagccgg99cagctcctcggccgggtg99gggtc99gggttccccag 1200
Db 1300 AACGAGTCTCTCATAGCCGGGCCGAGTTCTTGGCCGGGGTGGCGGGTTCGGGGTTCCCGCAG 1359

Qy 1201 gtaagtgaacctggcagccgaggtgtggtcctgcattacacagactgctgcaccccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCTATCCCGAG 1419

Qy 1261 gacccggcacgctgaggg99ccctgagcgtg99tgg99gacccacaatgtcgtgtgc 1320
Db 1420 GACCCGGCACGCTGAGGGAGGCCCTCAGCGATGTGTGTGGCGCACCAATAATGTGCTGTGC 1479

Qy 1321 ccggtggccagctgctggggcagctggtgccccaggtgccccgggtctcagcgcctacgctc 1380
Db 1480 CCGGTGGCCAGCTGGCTGGGGCAGCTGGCTGCCCCAGGGTCCCGGGGTCTACGCCCTACGTC 1539

Qy 1381 ttgaaacacagctgcttccacgctcctcgtgcccctgtggtgggtggtgccccacggtac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTCGCCCTGTGTGATGGGGGTGCCCCACCGGTAC 1599

Qy	1141	aacgagctctctcatcagccggygcccaggttctctggccggygtcggggtccgggttccccag	1201
Db	1300	AACGAGCTCTCTATCAGCCGGGCGGAGTTCTTGGCGGGGTGCGGGTTCGCCAG	1359
Qy	1201	gtaagtgaacctggcagccgagagctgtgtctctgcattacagagactggctgcatcccgag	1260
Db	1360	GTAAAGTGACTGCGACCGGAGGCTGTGTCTCTGCAATTACACAGACTGGCTGCATCCGAG	1419
Qy	1261	gacccggcacgctctgagggagccctgagcagtgatggggtggcgaccacaatgtcgtgtgc	1320
Db	1420	GACCCGGCAGCCTGAGGGAGGCCCTGAGCGATGTGTGGGGGACCACAATGTCTGTGC	1479
Qy	1321	cccggtggccagctggctggggtggcagctgctgcceagggtgcccggtgtctacgacctacgctc	1380
Db	1480	CCCGTGGCCAGCTGTGCTGGGGACTGTGGTGCCAGGTGCCCGGTCTACGCCCTACGTC	1539
Qy	1381	tttgaacacgctgtctccacgctctctctggccctctgtgatgggggtgccccacgagctac	1440
Db	1540	TTTGAACACGGTGTCTCCACGCTTCTTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC	1599
Qy	1441	gagatcgagttctcatctttggatcccccctggaccctctcgaaactacacggcgagagag	1500
Db	1600	GAGATCGAGTTTCATCTTTGGATCCCTCTGGACCCCTCTCGAAACTACACGGCAGAGAG	1659
Qy	1501	aaatcttgcgccagcgactgtcgatactggtggccaactttgccgcacaggggataccc	1560
Db	1660	AAATCTTGCGCCAGCGACTGTATGGCGATACTGGGCCAACTTTGGCCGCACAGGGGATCCC	1719
Qy	1561	aatgagcccgagaccceaggccccacaatgtgccccgtatacagcgcggggtctcagcag	1620
Db	1720	ATGTAGCCCCGAGACCCCCAAGGCCGCCACAATGGCCCCCGTACACGGCGGGGGCTCAGCAG	1779
Qy	1621	tacgttagtctgtgaacctgcggtcgctgtgaggttgcggtcggggggtgcgcgccagggctgc	1680
Db	1780	TACGTTAGTCTGGACCTGCGCGCGCTGGAGGTGCGCGCGGGGCTGCGCGCCACGCGCTGC	1839
Qy	1681	gcctctgtgaaccgctctctccccaaaattgtctcagcgccacc	1722
Db	1840	GCCTCTGTGAACGGCTTCTCCGCCAAATGTCTCAGCGGCAC	1881

RESULT 11

US-08-814-095-5
: Sequence 5, Application US/08814095
: Patent NO. 6025183
: GENERAL INFORMATION:
: APPLICANT: Soreq, Hermona
: APPLICANT: Zakut, Haim
: APPLICANT: Shani, Moshe
: TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
: TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: KOHN & ASSOCIATES
: STREET: 3050 N. 6025183rd Western Highway, Suite 410
: CITY: Farmington Hills
: STATE: Michigan
: COUNTRY: U.S.
: ZIP: 48334
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/814,095
: FILING DATE:
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Montgomery, Ilene N.
: REGISTRATION NUMBER: 38,972
: REFERENCE/DOCKET NUMBER: 2391.00066

```

Db 880 GGCATGACCTGCTGCTCCCGCCAGCGGGCCCTGTTCCACAGGGCGTGTGTCAGAGC 939
Qy 781 ggtgcccccaatggaccctgggacagctgggcatggagagggccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTGGGCCACGCTGGGCATGGGAGAGGCCGCTCGCAGGGCCACG 999
Qy 841 cagctggccccacctgtggtgtgtctccagggcgcaactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGGCCCACTTGTGGGTGTCTCTCCAGCGCGCACTGTGGGAATGACACAGAGTG 1059
Qy 901 gtgactgccttcgacacgacacagcagcaggtccctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCTGCTTCGACACAGCAGCAGCAGCAGGTCTGTGTGAACACAGATGGCACGCTGTG 1119
Qy 961 cctcaagaaagctgttcctcggttctctctgctgctgtgtagtaggagacttctctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTCCGGTTCCTCTCTGTGCTGTGGTAGATGGAGACTTCTCTCAGT 1179
Qy 1021 gacacccagagggccctcatcaacggggaggaacttccacggcctgcaggtgctggtggt 1080
Db 1180 GACACCCAGAGGGCCCTCATCAACGGGGAGACTTCCACGGCCTGCAGGTGCTGTGGTGGT 1239
Qy 1081 gtggtgaagatgagggctgtatttctgttttacggggcccgagccttcagaaagac 1140
Db 1240 GTGGTGAAGATGAGGGCTGTATTTCTGTGTTACGGGGGCCCGAGGCTTCAGCAAGAC 1299
Qy 1141 aacgagctctcatcagcggggcgaggttctctggcgggggtgcggggttcccccag 1200
Db 1300 AACGAGTCTCTCATCAGCGGGCGGAGTTCCTGGCGGGGTGGGGTTCGCCAG 1359
Qy 1201 gtaagtacctggcagcggaggtgtgtctgcatcaacagactgctgcacccag 1260
Db 1360 TTAAGTGAACCTGGCAGCGGAGGCTGTGTCTGTGCTGCAATACACAGACTGCTGATCCCGAG 1419
Qy 1261 gaccggcagcctgagagagccctgagcagatggtggcgacccaatgtcgtgtgc 1320
Db 1420 GACCGGACGCTGAGGAGGCGCTGAGCAGATGTGTGGGGGACCAACATGTCTGTGTGC 1479
Qy 1321 cccgtggcccgagctgggtggcgactggtcccgaggtgcccgggttctacgcttacgtc 1380
Db 1480 CCGTGGCGCCAGCTGGCTGGCGGAGTGGCTGCCAGGGTGCCTGGGTCTACGCTTACGCTC 1539
Qy 1381 ttgaaacacgtgtctcagcgtctctctggccctgtggtatgggggtgccccagggctac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTTGGCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599
Qy 1441 gagatcgagttcatctttgggatccccctggacccctctcgaactacacggcagagag 1500
Db 1600 GAGATCGAGTTCACTTTGGGATCCCCCTGGACCCCTCTGAACTACACGGCAGAGGAG 1659
Qy 1501 aaatcttcgccccgagactgatcgatactgtggcgaactttgcccgcacaggggagatccc 1560
Db 1660 AAATCTTTCGCCAGCGACTGATGATGATGAGGCACTTGGCGCACTTGGCCGACAGGGGATCCC 1719
Qy 1561 aatgagccccgagaccccaagcccccaaatggccccctgtacacggcgggggtcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGCCCCCAATGGCCCCGTACACGGCGGGGGTCTCAGCAG 1779
Qy 1621 tacgttagcttgacctgcgcccgtgaggtgcggcgggggtgcgcgccccagcctgc 1680
Db 1780 TACGTTAGCTTGGACCTCGGCCGCTGTGAGGTGGCGGGGGGTGCGCGCCCCAGGCCCTGC 1839
Qy 1681 gcctcttgaaacccgcttctcccaaaattgtctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCCCCAAATGCTCAGCGCCACC 1881

```

RESULT 12

US-08-814-095-7
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:

```

; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE
; DESCRIPTION: promoter, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon

```



```
; MAP POSITION: 3q26
; FEATURE: mat_peptide
; NAME/KEY: 160..1881
; LOCATION: 160..1881
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ECHE"
; OTHER INFORMATION: /note= "butyrylcholinesterase mature peptide"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 76..159
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: 1..2381
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..1884
; US-08-318-826A-9

Query Match      22.0%; Score 380; DB 2; Length 2381;
Best Local Similarity 55.2%; Pred. No. 1.5e-71;
Matches 790; Conservative 0; Mismatches 625; Indels 15; Gaps 2;

QY 146 gggggaattcgctgaagaccgccggggccctgtctctgtcttcttctgggcatccctttg 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 200 GAGGATGAACATGACAGTTTGGTGCGACGGTAACAGCCTTCTCTGGAATTCCTCATG 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 cggagccaccatggagccctgtcttctgccaccggagcccaagcagccttgtag 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 260 CACAGCCACCTCTGGTAGACTTCGATTCAAAAGCCACAGCTCTGACCAAGTGCTCG 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 ggggtgtagacactcaacactccagagtgctgtctaccatgatgtggacacctatacc 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 320 ATATTGGAATGCCAANAATATGCAAAATCTTGCTCAGAACATAGATCAAGTTTC 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 cagggttgaggccaccagagatgtggaaccacacccagctgagctgagcaggagcctgt 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 380 CAGGCTTCATCGATCAGAGATGTGGAACTGGAACCAACTGACCTCAGTGAAGACTGTTAT 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 acctcaactgtggacaccatccccccggcctacatcccccccccctgtcctgtctgga 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 ATCTAAATGATGGAATCCAGCACCTTAACCAAAAATATGCCA--CTGTATTGATGGA 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 tctatgggggtggtctctacagtggggccctctctcttggacgtgtacagtggccgtct 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 497 TTTATGGTGGTGTTCAAACTGGAACATCATCTTTACATGTTTATGATGGCAAGTTTC 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 tggtagagccagagagactgtctgtgtgtccatgaactaccgggtggagcctttggt 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 557 TGGCTCGGTTGAAAGAGTTATTTAGTGTCTCAATGAACATATAGGTGGGTGCCCTAGGAT 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 tcttgccctgcggggagcagagagcccgcccggaatgtgggtctcctggatcagagcc 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 617 TCTTAGCTTGCAGGAAATCTCTGAGGCTCCAGGGAAACATGGGTTATTGTATCAACAT 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 tggccctgcagtggtgtagggagacgtggcagccttccgggtgtagcccgacatcagta 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 677 TGGCTCTTCAGTGGGTTCAAAAAATATAGCAGCCTTTGGTGGAAATCTCTAAAGGTAA 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 686 cgtgtttgggagagcagcggagccgctcgtgtgggcatgacactgtgttcccccacca 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 737 CTCTCTTTGGAGAAATGTGAGGAGCAGCTTCAGTTAGCCTGCAATTTGCTTCTCTGGAA 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 746 gccggggccctgttccacagggccgtgctgctcagagcgtgtcccccaatgacccctggcca 805
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 797 GCATTCATGTTCCACGAGGCATCTCTCAAAAGTGATCTCTTTAATGCTCTTGGCGG 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 806 cgggtgggacagggagggccctcgagggccacgagcgtggccaccctgtgggtgtc 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 857 TAACATCTCTTTATGAAGCTAGGACAGAACGTTGAACTTAGCTTAATTT----- 905
```

```
QY 866 ctccagggcggcactggtgggaatgacacagagctggtgtagccttcggacacgaccag 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 906 -GACTGGTTGCTCTAGAGAGAAATGAGACTCAATAATAAGTCTCTAGAAATAAAGATC 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 926 cgcaggtcctggtgaaccacgaatgscagctgctcctcaagaagcgtcttcggtct 985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 965 CCCAAGAAATCTCTGAATGAAGCATTTGTTGTCCTTATGGGACTCTCTTGTCTAGTAA 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 986 ccttgctgctggtgtagatgagacttctcagtcagtcacaccccgagggccctcatcaag 1045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1025 ACTTTGGTCCGACCGTGGATGGTGAATTTCTACTCAGATGCCAGACATATTACTTTGAAC 1084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1046 cgggagacttccacggcctgcagggtgctggtggtggtggtggtggtggtggtggt 1105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1085 TTGGACAATTTAAAAAACCCAGATTTTGGTGGGTCTTAATAAAGATGAAGGACAGCTT 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1106 ttctggtttacggggcccgagccttcagcaagacacagctctctcctcagcgggccc 1165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1145 TTTTAGTCTATGGTCTCTGCTTCCGCTTCAGCAAAAGATTAACAATAGTATCAATACTAGAAAG 1204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1166 agttctgcccgggtgctgctggtggtggtggtggtggtggtggtggtggtggtggt 1225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1205 AATTTTCAGGAAGGTTTAAAAATATTTTTCAGGAGCTGAGTGAGTTTGGAAAGGAATCCA 1264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1226 tggctcgtcattacacagactggctgcacccagagaccccgccagcctgagggagccc 1285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1265 TCCTTTTTCATTACACAGACTGGGTAGATGATCAGAGACCTGAAAACTACCGTGAAGCCT 1324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1286 ttagcagatggtggcgaccacaatgtgctgtgccccctggccagctggctggcgac 1345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1325 TGGTGATGTTGTTGGGATTAATAATTCATATGCCCTGCCTTGGAGTTTACCACAAAGT 1384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1346 tggctgcccaggtgcccgggtctacgctcgtctttgaacacccgtgtctccacgctct 1405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1385 TCTCAGAATGGGAAATAATGCTTTTCTACTACTTTTGAACACCGATCTCCCAAACTTC 1444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1406 cctggccctgtggtggtgggtggtggcccggtacagatcgagttcatcttctgggagtc 1465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1445 CGTGGCCAGAAATGGATGGGAGTGCATGCATGGCTATGAAATTTGAAATTTGTTTGTTC 1504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1466 cctggagccctctcgaactacacgagcagagagagaaatctctcgcagcagctgctgc 1525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1505 CTCTGGAAGAGAGATATATTACAAAAGCCGAGGAATTTTGTAGTAGATTCATAGTGA 1564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1526 gatactggccacttggcccgacagggggtaccccaatgagcccccgagac 1575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1565 AACGGTGGCAAAATTTTGCAAAATATGGGAATCCAAATGAGACTCAGAAC 1614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-08-318-826A-8
; Sequence 8, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:04:30 ; Search time 9324.87 Seconds
(without alignments)
12942.093 Million cell updates/sec

Title: US-09-810-861b-3
Perfect score: 5767
Sequence: 1 agctgcagctgcaggc.....ctatgaccatgattacgccca 5767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb_ba:**
- 2: gb_htg:**
- 3: gb_in:**
- 4: gb_om:**
- 5: gb_ov:**
- 6: gb_pat:**
- 7: gb_ph:**
- 8: gb_pl:**
- 9: gb_pr:**
- 10: gb_ro:**
- 11: gb_sts:**
- 12: gb_sy:**
- 13: gb_un:**
- 14: gb_vi:**
- 15: em_ba:**
- 16: em_fun:**
- 17: em_hum:**
- 18: em_in:**
- 19: em_mu:**
- 20: em_om:**
- 21: em_or:**
- 22: em_ov:**
- 23: em_pat:**
- 24: em_ph:**
- 25: em_pl:**
- 26: em_ro:**
- 27: em_sts:**
- 28: em_un:**
- 29: em_vi:**
- 30: em_htg_hum:**
- 31: em_htg_inv:**
- 32: em_htg_other:**
- 33: em_htg_inv:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match Length	DB ID	Description
------------	--------------------	-------	-------------

1	5767	100.0	5767	6	AX275254	AX275254 Sequence
2	2637	45.7	3858	6	AX114873	AX114873 Sequence
3	2637	45.7	4473	12	UCU89964	U89964 Cloning vec
4	2637	45.7	4960	6	AX191663	AX191663 Sequence
5	2637	45.7	5711	6	AX114861	AX114861 Sequence
6	2637	45.7	7332	6	AX191664	AX191664 Sequence
7	2637	45.7	8062	6	AX114872	AX114872 Sequence
8	2637	45.7	8153	6	AX114871	AX114871 Sequence
9	2636	45.7	2746	12	CVPCEM42	X65305 Cloning vec
10	2635.4	45.7	4392	12	CVU47294	U47294 Cloning vec
11	2635.4	45.7	4518	12	XXP35SGFP	U28417 Cloning vec
12	2635.4	45.7	4519	12	U02456	U02456 Cloning vec
13	2635	45.7	4674	12	XXU02437	U02437 Cloning vec
14	2634.4	45.7	5027	12	AF276982	D276982 Integrati
15	2633.8	45.7	5647	1	BACKNB	D29979 Bacillus st
16	2633.8	45.7	11612	12	CVU80929	U80929 Cloning vec
17	2633.8	45.7	16297	12	CVU75991	U75991 Cloning vec
18	2633.8	45.7	16437	12	AF123770	AF123770 Shuttle v
19	2633.8	45.7	18772	12	AF133437	AF133437 Cloning v
20	2633.8	45.7	19500	12	CVU75992	U75992 Cloning vec
21	2633.4	45.7	2743	12	CVPCEM32	X65304 Cloning vec
22	2632.8	45.7	7387	6	AR176236	AR176236 Sequence
23	2632.8	45.7	7387	6	AX003207	AX003207 Sequence
24	2631.8	45.6	4283	6	I63120	I63120 Sequence 3
25	2631.8	45.6	4283	6	I85496	I85496 Sequence 1
26	2630.2	45.6	3394	12	AF062083	AF062083 Cloning v
27	2630.2	45.6	3394	12	AF083409	AF083409 Cloning v
28	2630.2	45.6	3604	12	AF062078	AF062078 Cloning v
29	2630.2	45.6	3604	12	AF083407	AF083407 Cloning v
30	2630.2	45.6	3617	12	AF062079	AF062079 Cloning v
31	2630.2	45.6	3714	12	AF083408	AF083408 Cloning v
32	2630.2	45.6	3946	12	AF134573	AF134573 Cloning v
33	2630.2	45.6	4026	12	AF062080	AF062080 Cloning v
34	2630.2	45.6	4154	12	AF062082	AF062082 Cloning v
35	2630.2	45.6	4503	12	AF062081	AF062081 Cloning v
36	2630.2	45.6	4795	12	AF134572	AF134572 Cloning v
37	2630	45.6	3796	6	E49323	E49323 Infectious
38	2630	45.6	5452	12	U02454	U02454 Cloning vec
39	2630	45.6	6541	12	XXU02433	U02433 Cloning vec
40	2630	45.6	6894	12	XXU02435	U02435 Cloning vec
41	2630	45.6	7069	12	U02442	U02442 Cloning vec
42	2630	45.6	7164	6	AX195206	AX195206 Sequence
43	2630	45.6	7164	12	U02451	U02451 Cloning vec
44	2630	45.6	7481	12	XXU02438	U02438 Cloning vec
45	2628.4	45.6	15528	6	A93016	A93016 Sequence 4

ALIGNMENTS

RESULT	1					
AX275254						
LOCUS	AX275254	5767 bp	DNA	linear	PAT 29-OCT-2001	
DEFINITION	Sequence 3 from Patent WO0171014.					
ACCESSION	AX275254					
VERSION	AX275254.1	GI:16547674				
KEYWORDS	synthetic construct.					
SOURCE	synthetic construct					
ORGANISM	artificial sequence.					
REFERENCE	1 (sites)					
AUTHORS	Mor, T., Soreq, H., Arntzen, C. and Mason, H.					
TITLE	Expression of recombinant human acetylcholinesterase in transgenic plants					
JOURNAL	Patent: WO 0171014-A 3 27-SEP-2001; BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor, Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)					
FEATURES	Location/Qualifiers					
source	1. 5767					
	/organism="synthetic construct"					
	/db_xref="taxon:32630"					
	/note="plasmid vector pTM034"					

BASE COUNT		1390 a	1521 c	1495 g	1361 t		
ORIGIN							
Query Match		100.0%;		Score 5767;		DB 6; Length 5767;	
Best Local Similarity		100.0%;		Pred. No. 0;			
Matches 5767;		Conservative 0;		Mismatches 0;		Indels 0; Gaps 0;	
QY	1	agcttgatcgctcgaggtcaacatggtggagcagcacactctcgtctactccaagaata	60				
DB	1	AGCTTGCAATGCTGCAAGTCAACATGCTGGAGCAGCACACTCTCGTCTACTCCAAGAA	60				
QY	61	tcaaaagatacagtcctcagaagaccagagggctattagagactttcaacaaagggtaatat	120				
DB	61	TCAAAAGATACAGTCTCAGAAAGACAGAGGGCTATTGAGACTTTTCAACAAAGGGTAATAT	120				
QY	121	cgggaacactcctcgattccattcgccagctatctgtcacttcacgaagagacagtag	180				
DB	121	CGGGAACCTCTCGGATTCCATTGCCCCAGCTATCTGTCACTTCATAAGGACACAGTAG	180				
QY	181	aaaaggaagatggcttctcaaaatgccatcattgcgataaaagaaaggctatcgttcaag	240				
DB	181	AAAAGGAAGATGGCTTCTACAAATGCCATCATTTGCGATAAAGGAAAGGCTATCGTTCAAG	240				
QY	241	aatgccttacgacaagtagtggccccaaagatggaaccccccaacgaggaacatcgtagaa	300				
DB	241	AATGCCCTCTACCGACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAACATCGTGGAA	300				
QY	301	aaagaagacgttccaaaccgctcttcaagcaagtgtgattgatgaacttttcaaca	360				
DB	301	AAAGAAGACGTTCCAAACCACGCTTTCMAAGCAAGTGGATTGATGTGATTAATTTTCAACA	360				
QY	361	aagggtaatatcgggaaacctcctcgattccattgccagctatctgtcactttcatcga	420				
DB	361	AAGGGTAATATCGGAAACCTCTCGGATTCCATTGCCAGCTATCTGTCACTTTCATCGA	420				
QY	421	aaggacagtagaaaagaagtagtggcttctacaaaatgcataatcgatgaagaaagc	480				
DB	421	AAGGACAGTAGAAAAGAAAGATGGCTTCTACAAATGCCATCATTTGCGATAAAGGAAAGGC	480				
QY	481	tatcgttcaagaatgcctctacgcagctgggtcccaagaatggaacccccacacagagga	540				
DB	481	TATCGTTCAAGATGCTCTACCGACAGTGGTCCCAAAAGATGGACCCCCACCCACGAGGA	540				
QY	541	acatcgtggaanaaagacgtttccaaacacagctcttcaagaagtggtattgatgtgata	600				
DB	541	ACATCTGGAAAAAAGAACGTTTCCAAACACAGCTTCTCAAAGCAAGTGGATTGATGTGATA	600				
QY	601	tctcaactgaagtaagggatgacgaacaatcccactatctctcogaagaccttctctca	660				
DB	601	TCTCCACTGACGTAAGGGATGACGCAATGCCACTATCTCTTGGCAAGACCCCTTCTCTCA	660				
QY	661	tataaggaagttcatcttattggagaggaacctcgagaatttaattctcaacacaacatat	720				
DB	661	TATAAGGAAGTTCAATTATTGGAGAGACCTCGAGAATTAAATCTCAACACACATAT	720				
QY	721	acaaacaacaaatcgaacaaatcaagcaattctctactctatttgagcaaatttaataca	780				
DB	721	ACAAAACAACAAATCTCAAGCAATCAAGCAATCTACTTCTATTTCAGCAAAATTTAAATCA	780				
QY	781	tctctttaaagcaaaagaattttctgaaaaattttcaccatttaacgaacgatagccatg	840				
DB	781	TTTCTTTTAAAGCAAAAGCAATTTTCTGAAAATTTTTCACCAATTTAGCAACGATAGCCATG	840				
QY	841	gctcccccgagtgctgtgcacagccttccctggcttccccactctctctccctc	900				
DB	841	GCTCCCCCGAGTGTGTGTGCACAGCCCTTCCCTGGCTTCCCCTCTCTCTCTCTCTCTC	900				
QY	901	ctctggctcctggtaggagtagtgggggctgagggcgaggagatgcagagctgctgggtg	960				
DB	901	CTCTGGCTCTCTGGGTGGAGAGTGGGGGCTGAGGGCGGGAGGATGCAGAGCTGCTGGTG	960				
QY	961	acgggtgcgtgggggccggtgcggggcattcgcctgaagacccccggggccctgtctct	1020				

DB	961	ACGGTGCgtggggccggctgcggggcattgcctgaagaccggggccctgtctct	1020				
QY	1021	gcttctctgggaatccccctttgcggagccaccatgggagcccgctgccttctgccaacg	1080				
DB	1021	GCTTCTCTGGGATACCCCTTTCGGAGACCAACCATGGGACCCCGCTGCTTCTGCCACCG	1080				
QY	1081	gagcccaagcagccttggtcaggggtgtagacgctacaaccttccagagtgctctgtac	1140				
DB	1081	GAGCCCAAGCAGCCTTGGTCAGGGTGGTAGACGCTTACAACCTTCAGAGTGTCTGCTAC	1140				
QY	1141	caatatgtggacaacctataaccagggttttgagggcaccgagatgtggaacccccaccgt	1200				
DB	1141	CAATATGTGGACACCCCTATACCCAGGTTTGGGGCACCGAGATGTGAACCCCAACCGT	1200				
QY	1201	gaagtgcgaggaactgcctgtacctcaactgcagctggacaccatcccccgccctacatcc	1260				
DB	1201	GAGCTAGCGAGGACTGCTGTACCTCAAGCTGTGACACCATACCCCGGCTTACATCC	1260				
QY	1261	ccacccctgtctcgtctggatctatgggggttgggttctctacagtgggggcctcctctg	1320				
DB	1261	CCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTTCTCTCTTG	1320				
QY	1321	gaagtgcagatggcgcttcttgggtacaggccgagaggaactgtgctgtgtccatgaac	1380				
DB	1321	GAGGTACGATGGCGCTTCTTGGTACAGCGGAGAGGACTGTGTGTGTCCATGAAC	1380				
QY	1381	taccgggtggagccttttgcttctcctcggccctgcggagccgagagggcccgggcaat	1440				
DB	1381	TACCGGGTGGAGCCTTGGCTTCTTGGCTTCCCTGCCGGGAGCGAGGCCCCGGGCAAT	1440				
QY	1441	gtgggtctcctggatcagaggctggccctgcagtggtgcaggagaacgttggcagccttc	1500				
DB	1441	GTGGGTCTCTGGATCAGAGCTGGCCCTGCAGTGGGTGCAGGAAACGTGGCAGCTTC	1500				
QY	1501	gggggtgacccagacatcagtgacgtgtttggggagagcgcgggagcccgctcgtgggc	1560				
DB	1501	GGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGGGCCCTCGGTGGGC	1560				
QY	1561	atgcacctgctgtccccgcggcgccgtgttccacagggccgtgctgcagagcggt	1620				
DB	1561	ATGCACCTGCTGTCCCCGCCACGCCGGGCTGTTCACAGGGCCCTGTGTGCAGAGGGT	1620				
QY	1621	gcccccaatggaccttggccacgggtgggaatgggagagcccgctgcagggccacgag	1680				
DB	1621	GCCCCAAATGGAGCCCTGGGCCACAGGTGGGCTGGGAGAGCGCCGTGCAGGGCCACGAG	1680				
QY	1681	ctggccacactgtgggctgtctccagcgcgcaactgttgggaatgacacagagctggta	1740				
DB	1681	CTGGCCACCTTGTGGGCTGTCTCCAGGCGGCACCTGGTGGGAATGACACAGAGCTGGTA	1740				
QY	1741	gcctgcttcggacacgacgacgagcgaggttctggttgaaacacgaaatggcagctgctg	1800				
DB	1741	GCTGCTTCCGACACACACAGCGCAGGTCTTGTGTGAACCAACCAAGTGCAGCTGTGCT	1800				
QY	1801	caagaagcgtcttcgggttctcctctgctgtgctgtgtagatggagacttctcagtgac	1860				
DB	1801	CAAGAAGCGCTCTTCCGGTCTCTTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTG	1860				
QY	1861	acccagagccctcatcaacgagagacttccagcgctgcagctgctggtggtgtg	1920				
DB	1861	ACCCAGAGGCCCTCATACCGGGGAGACTTCCAGGGCTGCAGGTGTGTGTGTGTGTGTG	1920				
QY	1921	gtgaagatgagggctcgtatttctggttttacggggccccaggcttcagcaagaac	1980				
DB	1921	GTGAAGATGAGGGCTCGTATTTCCTGGTTTACGGGGCCCCAGGCTTCAGCAAGACAAAC	1980				
QY	1981	gagctctcatcagccggggccaggttctcggccgggtgctgggtcggggttccccaggta	2040				
DB	1981	GAGTCTCTCATCAGCCGGGCGAGTTCTTGGCCGGGTGCGGGTTCGGGTTCGCCAGGTA	2040				
QY	2041	agtgaactggagcagaggtgtgtctctgattacagagactgctgcatccccggagac	2100				

Db 2041 AGTGACCTGGCAGCGGAGCTGGTCTGCTGCATTTACAGACACTGGCTGCATCCCGAGGAC 2100
Qy 2101 ccggcaacgctgaggggcccctgagcga tgtggtggcgccacacaaatgtcgtgtcccc 2160
Db 2101 CCGCAGCCCTGAGGAGGCGCTGAGCGATGTGTGGCGGACCAACAATGTCTGTGCCCC 2160
Qy 2161 gtggccagctggctggcgacgtgctgccaggggtgcccgggtctcagcctaagctttt 2220
Db 2161 GTGGCCAGCTGGCTGGCGGACCTGGCTGCCCCAGGGTGCCCGGGTCTACGGCTACGCTTT 2220
Qy 2221 gaacacgctgtctcaacgctctctggccctctggtggtgggggtgcccacaggtacag 2280
Db 2221 GAACACCGTGTCTCCACGCTCTCTGTGCCCCCTGTGGATGGGGTGCCCCACGGCTACGAG 2280
Qy 2281 atcgagttcatcttttggga tccccctggacccctctctcgaactacacgagagagaaa 2340
Db 2281 ATCGAGTTCAATCTTTGGGATCCCGCTGCAGCCCTCTCGAAACTACAGCGCAGAGAGAAA 2340
Qy 2341 atcttcgcccagcagctgactgactgagcgaactttgcccgacagggggtcccaat 2400
Db 2341 ATCTTCCGCCAGCGACTGTATGCGATATGGGCCAACTTTGCCCGCACAGGGGATCCCAAT 2400
Qy 2401 gagcccgagaccccaaggcccacaaatggcccctgtacacggcgggggtcagcagttac 2460
Db 2401 GAGCCCCGAGNCCCAAGGCCCAACATGGCCCCCTGTACAGGGCGGGGCTCACAGCTAC 2460
Qy 2461 gttagtgtgacctgycgcgtgaggtggcggtggcggtggcggtggcggtggcggtggcg 2520
Db 2461 GTTAACTGGACCTGGCGGCGCTGGAGGTGGCGGGGGCTGGCGGCCACGGCTGGGCC 2520
Qy 2521 ttctggaaacgcttctcccaaaattctcagcgctaccctgatataggtagcagctctctc 2580
Db 2521 TTCTGGAACCGCTTCTCCCAAAATGCTCAGCGCTACCTGTATAGGTACCGAGCTCTCTC 2580
Qy 2581 aacaactagtagagttgtctctctctataatgaataaggtatgctgtatcactat 2640
Db 2581 AACAACTAGCTAGATTGCTCTCTATCTATATGTAATAAGGTATGCTGATATGCACTAT 2640
Qy 2641 tcaaataggagcattagctatgtttgtaattgtcaactttatgttatgtgggaagtcaacc 2700
Db 2641 TCAAAATAGGACATATGAGTCTTTTGTAAATGTACATTTATGTTATGTGGTAAAGTCACC 2700
Qy 2701 taagacactcacactacgtctgtgtctcttaccgctttaaataactcttgcct 2760
Db 2701 TAAGACACTCCACCTACTAGTTGTGTCTCTTACC GGCTTAAATAATCTTCTGGCCCT 2760
Qy 2761 tgttccattattactaatatcccccttctcactaaagaataattgtttatcatlaagtat 2820
Db 2761 TGTTCCATATTTACTAATATATCCCTTCTTCACCTAAAGAAATTTGTTATCATTAAGTAT 2820
Qy 2821 tagctttagaacatataggtctttaaattgggtaggttttcacaaattcaactaatataa 2880
Db 2821 TAGTCTTTAGAACATATGAGTCTTTAAATTTGGGTAGGTTTACAAATTAACATAATAA 2880
Qy 2881 atgtcaataaatcacatggttaacaaatcgaaaaatcgacgtctctattgacga 2940
Db 2881 ATGTCTATAAATCCAGCTGGTGTAAACAAATGCAGAAATTCAGCTGTCTATTGGACCGA 2940
Qy 2941 cagttgctattaaataatgggccaccatagtagctgacaaaataaaattacctgacaaca 3000
Db 2941 CAGTTGCTATTAAATATATGGGCCACCATAGTAGCTGACAAATAAATTAACCTGACRACA 3000
Qy 3001 tcgtttcaactaaa taacaaacacaaaagggtgagctgtttccaggggcatttttgttaata 3060
Db 3001 TCGTTTCACTAAATAACAAACACAAAAGGAGTGCATTTTCCAGGGCATTTTGTAAATA 3060
Qy 3061 aaaaacagttaaaggagtgcaatagaataataggggtgtggaatagtattgagca 3120
Db 3061 AAAAACAGTTAAAAAGGAGTGCATAGAAATATAGGGGTGTGGAAATAGTGAATTTGAGCA 3120
Qy 3121 cgtctgaagcgaattcaactgctgtctgttttaaacgctgtgactggggaacacccctggc 3180
Db 3121 CGTCTTGAGCGAATTCACCTGGCGCTCGTTTTACACACCTCGTACTGGGAAAACCCCTGGC 3180

Qy 3181 gttaccaacttaatgccttgcagcacatccccctttgcagcagctggcgttaa tagcgaa 3240
Db 3181 GTTACCACAACTTAATCGCCTTGCAGACATCCCCCTTTGCGCAGCTGGCGTAATAGCGAA 3240
Qy 3241 gagcccgacacccagctcccttcccaagttgcagacgtgaatgcgaatgagcgtg 3300
Db 3241 GAGCCCCGACCGATGCGCCCTCCCAACAGTTGCGAGCCTGAATGGCGAATGGCGCCTG 3300
Qy 3301 atgcggtattttcccttaagcatctgtgcggtatttcaacacccgcaatatggtgcactctc 3360
Db 3301 ATGCGGTATTTCTCTTACGCATCTGTGCGGTATTTTCACACCCGCATATGTCACATCTC 3360
Qy 3361 agtacaactgctgctgtagcgcagtagttaagcagccccgcacccccgcacccccgcct 3420
Db 3361 AGTAAATCTCTGTATGCGCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCT 3420
Qy 3421 gagcgccttgacgggctgtctgctcccgcatccgcttacagacaagctgtgacccctc 3480
Db 3421 GACCGCCCTGACCGGCTTGTCTGCTCCCGCANTCCGGTTACAGACAAGCTGTGACCCGTC 3480
Qy 3481 tccgggagctgcatgtgtcagaggttttccacgtcatcacccgaaacgcgcgagacgaag 3540
Db 3481 TCCGGAGCTGCATGTGTACAGAGTTTTTCACCGTCATCACCGAAACGCGCAGACGAAAG 3540
Qy 3541 ggcctcgtgatacagcctatttttaggttaagtcatgataaataaattggttcttagacg 3600
Db 3541 GGCCCTCGTATACGCCATATTTTATAGTTAATGTATCATGATAAATAAGTGTCTTAGACG 3600
Qy 3601 tcagttggcacttttcgggaaatgtgcggaacccctatttatttttttctaaata 3660
Db 3601 TCAGTTGGCACTTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTTAAATA 3660
Qy 3661 cattcaaatatgtagtattcaacatttccgctgctgcacctattcccttttttgcggca 3720
Db 3661 CATTCAAATATGATCCGCTCATGAGACAATAACCTGATTAATGCTTCAATAATATGA 3720
Qy 3721 aaaaaggagagtagattcaacatttccgctgctgcacctattcccttttttgcggca 3780
Db 3721 AAAAGGAAGAGTATGAGTATTAACAATTTCCGTGTGCGCCCTATTTCCTTTTTCGGCA 3780
Qy 3781 tttgcctctcgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagat 3840
Db 3781 TTTTGCCTTCTGTTTTTGTCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGAT 3840
Qy 3841 cagttgggtgcacagtggtgtacatcgaaactggtatcacaacggtgaagatcccttag 3900
Db 3841 CAGTTGGGTGCACGAGTGGTGTACATCGAAGTGGATCTCAACAGCGGTAAAGATCCTTGAG 3900
Qy 3901 agtttgcgcccgaagaacgcttttccaaatgtagagcacttttaaagttctcgtatgtgc 3960
Db 3901 AGTTTTCGCCCGAAGAACGTTTTTCCAATGATGAGCACCTTTTAAAGTTTCTGCTATGTGC 3960
Qy 3961 gcgttattatccgtattgacgcgggcaagagaactcgtgcgcgcacacactatct 4020
Db 3961 CGGTATTATCCGTTATTGACGCCGGGCAAGCAACCTCGGTGCGCGCATACACTATTCT 4020
Qy 4021 cagaatgactgtgtgtagtactcaccagctcacagaagaagcatcttcacgtaggcatgaca 4080
Db 4021 CAGAATGACTTGGTTGAGTACTCACACAGTACAGAAAAGCATCTTACGGATGGCATGACA 4080
Qy 4081 gtaagagaattatgtagtgcgtgcaataaccatgagtgaatacactgcggccaacttactt 4140
Db 4081 GTAAGAGAAATATGCAGTGCTGCCATAACCATGAGTGAATAACACTGCGGCCAACTTACTT 4140
Qy 4141 ctgacaacgatcgagagacgaagagctaacccgttttttgacaacataggggatcat 4200
Db 4141 CTGACACGATCGGAGGACCGAAGGAGTAAACCGTTTTTTTGCACAACTATGGGGATCAT 4200
Qy 4201 gtaactgccttgcgttgggaacccgagctgaatgaagccatataccaaacgacgagcgt 4260
Db 4201 GTAACCTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGCGT 4260

Qy	4261	gacacacgagtgcctgtagcaatgccaacacgcttgcgcaacatttaactcggcgaaacta	4320
Db	4261	GACACCACGATGCCGTGTAGCAATTGGCAACACAGCTTGGCCAACTATTTAACCTGGCGAACTA	4320
Qy	4321	cttactctagcttcccggcgaacaattaaatagactcgatggagcgcgataaagtctgcagga	4380
Db	4321	CTTACTCTAGCTTCCCGGCACACAAATTAATAGACTCGATGGAGCGGAGTAAGATTTGTCAGGA	4380
Qy	4381	ccactctgcgctcggcccttcggctggctggtttattgtctgataaactctggagccggt	4440
Db	4381	CCACTCTGCGCTCGGCCCTTCGGCTGCCTGGTTTATTGCTGATAAATCTGGAGCCGGT	4440
Qy	4441	gagcgtgggtctcgcggtatcatctgcagcactcggggccagatggttaagccctcccgatc	4500
Db	4441	GAGCGTGGGTCTCGCGTATCATTTAGCTCAGCACTTGGGGCCAGATGGTAAGCCCTCCCGTATC	4500
Qy	4501	gtacttatctacacgacggggagtcaggcaactatggatgaacgaatatagacagatcgct	4560
Db	4501	GTACTTATCTTACACGACGGGGAGTCAGGCAACTATTGGATGAACCAAAATAGACAGATCGCT	4560
Qy	4561	gagataggctgcctcactgattaaagcatgtgtaactctgcagaccagtttactcaatatata	4620
Db	4561	GAGATAGTGCCTCACTGATTAAAGCATTTGGTAACCTGTGCAGACCAGTTTACTCATATATA	4620
Qy	4621	ctttagattgattaaacttcaatttttaatttaaaggatctagggtgaagatcctctttt	4680
Db	4621	CTTTAGATTGATTAAACTTCAATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTT	4680
Qy	4681	gataatcctatgacaaaatcccttaacgtgaagtttttcgttccactgagcgctcagaacccc	4740
Db	4681	GATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGGTCAAGACCC	4740
Qy	4741	gtagaaagatcaaaagatctctcttgagatcccttttctctgcgcgtaaactctgctgcttg	4800
Db	4741	GTAGAAAGATCAAAGATCTCTTGTGAGATCCCTTTTTTCTGCGCGTAATCTGCTGCTTG	4800
Qy	4801	caacaaaaaaaccacgcgtacacgogtgggttcttgcgggatcaagagctaccacact	4860
Db	4801	CAACAAAAAAACACCGCTACACGCGTGGTTTGTTCGCGGATCAAGAGCTACCAACT	4860
Qy	4861	cttttccgaagtgtaactggcttcacgagcgcagataccaaatactgtcccttctagt	4920
Db	4861	CTTTTTCCGAAGGTAACTGGCTTCAGAGAGCGCAGATACCAATAATCTGCTTCTTAGTG	4920
Qy	4921	tagcgttagttaggccaccacttcaagaactctgtagcacccgctcacaacctcgtcgt	4980
Db	4921	TAGCGGTAGTTAGGCCACCACCTTCAAGAACCTCTGTAGCACCGCCTCATACCTCGCTCTG	4980
Qy	4981	ctaatcctgttacagtggtgctgctccagtgggcgataagctgcttaccggggttggac	5040
Db	4981	CTAATCTGTTTACAGTGGCTGTGCCAGTGGCAGTAAGTCGTCTTACCGGTTGGAC	5040
Qy	5041	tcaagacgatagttaccggataaggcgcagcggctcggtctgaacggggggttcgtgcaca	5100
Db	5041	TCAAGACGATAGTTTACCGGATTAAGCGCAGCGGTTCGGCTGAAACGGGGGTTCTGTGCACA	5100
Qy	5101	cagccacgcttgagcgaaacgacctacacgaactgagatacctcacagcgtgagctatga	5160
Db	5101	CAGCCACGCTTGAGCGAACAGCCTTACCCGNACTGAGTACCTTACAGCGTGTAGCTATGA	5160
Qy	5161	gaaagcgccacgcttccccgaagggaagcggacaggtataccggtaagcggcagggctc	5220
Db	5161	GAAAGGCCACGCTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGCGAGGTC	5220
Qy	5221	ggaacgaggagcgcacgaggagcttccagggggaaacgcctggtaictttatagtcct	5280
Db	5221	GGAAACGAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCGCTGGTATCTTTTATAGTCCT	5280
Qy	5281	gtcggggttctgcacactctgacttgagcgtcgatttttctgatctcgtcagggggcgg	5340
Db	5281	GTGGGGTTTCGCCACCTCTGACTTGAGCGCTCGATTTTTTGATGATGCTCTGTCAGGGGGCGG	5340
Qy	5341	agcctatgaaaaaacccagcaacgcggcctttttacgggttctctgcgcttttctgctgcct	5400

Db	5341	ACGCTATGGA	AAAAAGCCAGCAACGCGCCCTTTTACGGTTCC	TGGGCTTTTGTGGCCT	5400
Qy	5401	tttgctcacatgttcttcctcgcttatccctcgattctctggtataaacctgattaccgccc	5460		
Db	5401	TTTGCTCACATGTTCTTTCCTGGCTTATCCCTGATTTCTGTGGATACCGGTATTACCGCC	5460		
Qy	5461	tttgagtgtgtatcacgctcgccgacgcgaacgaccgagcgcagtgctgagtcagtcgagc	5520		
Db	5461	TTTGAGTGAGCTGATACCGCTCGCGCAGCGAAGCACGAGCGACGAGTCAGTGAGC	5520		
Qy	5521	gaggaagcgggaagcgcgcgaataacgcaaacgcgcctctcccgcgcttggtggcgaattcat	5580		
Db	5521	GAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAT	5580		
Qy	5581	taatgcagctgacgcacaggtttcccgactggaaagcgggcgagtcgacgcgaacgcgaatt	5640		
Db	5581	TAAATGCAGCTGGCAGCAGCAGGTTTCCCGACTGGAAAGGGGCGAGTGAGCGAAGCAAT	5640		
Qy	5641	aatgtgagttagctcaactcattagcaaccgagctttacactttatgcttcggtcgctcgt	5700		
Db	5641	AATGTGAGTTAGCTCACTCAITAGCACCCCGAGCTTTACACTTTATGCTTCCGGCTCGT	5700		
Qy	5701	atgttgtgtgaattgtgagcggataacaaattcacacaggaagaaacagctatgacctgat	5760		
Db	5701	ATGTTGTGTGAATTGTGAGCGGATAACAAATTTACACAGGAACAGCTATGACCATGAT	5760		
Qy	5761	tacgccca	5767		
Db	5761	TACGCCCA	5767		
RESULT 2					
LOCUS	AX114873	AX114873	3858 bp	DNA	linear
DEFINITION	Sequence 33 from Patent WO0129208.				
ACCESSION	AX114873				
VERSION	AX114873.1	GI:14031815			
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	artificial sequence				
REFERENCE	1 (bases 1 to 3858)				
AUTHORS	Kuehn, R., von Melchener, H. and Altschmied, J.				
TITLE	Conditional gene trapping construct for the disruption of genes				
JOURNAL	Patent: WO 0129208-A 33 26-APR-2001;				
FEATURES	ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)				
source	Location/Qualifiers				
	1. 3858				
	/organism="synthetic construct"				
	/db_xref="taxon.32630"				
	/note="vector, PRK50"				
BASE COUNT	943 a	959 c	994 g	962 t	
ORIGIN					
Query Match					
Best Local Similarity	45.7%	Score 2637;	DB 6;	Length 3058;	
Matches 2637;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	3131	cgaattcacctggcgcgtgcttttacacgctgctgactgggaacccctggcgcttaccacaac	3190		
Db	1187	CGAATTCACTGGCCGTCGTTTACAAGCTCGTGAACCTGGGAACCCCTGGGCTTACCCNAC	1246		
Qy	3191	ttaatgccttgacgacatccccctttccagctggcgtaataagcgaagagcccgca	3250		
Db	1247	TTAATCGCTTTGCAGACATCCCTCTTCGCCAGCTGGCGTAAATAGCGAAGAGCCGCA	1306		
Qy	3251	ccgatacccttcccacagttgacgagcctgaatgagcgaatggcgctgacgctgatt	3310		
Db	1307	CCGATCGCCCTTCCCAACAGTTGCGAGCTGAATGCGAATGCGCGCTGATCGGTATT	1366		
Qy	3311	tctccttaacgcatctgtgcgqattttcacacgcgcataatgctacactctcagTacaactct	3370		

|||||
Db 1367 TTCCTTACGCATCTGTGCGGTATTTTACACCGCATATGTGCACTCTCAGTACAACT 1426
Qy 3371 gctctgatgcgcgatagtaagcagccccgacaccccgcccaaccccgctgacgcgcct 3430
Db 1427 CTTCTGATGCCGCATAGTTAAGCCAGCGCCGACACCCGCCAACCCGCTGACGGCCCT 1486
Qy 3431 gacgggcttgctgctcccgcatcccgcttacagacaagctgtgacgcctcctccggagct 3490
Db 1487 GACGGGCTTGCTGCTCCCGGCATCCGGTTACAGACAAGCTGTGACCCGTCCTCCGGAGCT 1546
Qy 3491 gcatgtgtcagaggttttcacgctcatcacccgaacgcgcgagacgaagggcctcgtga 3550
Db 1547 GCATGTGTCACAGGTTTTCACCGTCATCACCGAAACGGCGGAGAGAAAGGGCCTCGTGA 1606
Qy 3551 taccgctattttatagtttaagtcatgataataatggttttcttagacgtcaggtggca 3610
Db 1607 TACGCCATATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGCA 1666
Qy 3611 cttttcggggaaatgtgcgcggaacccctatttggttatttttcttaataatacatcaata 3670
Db 1667 CTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAATAATCATTCAAATA 1726
Qy 3671 tgtatccgctcatgagacaaataaaccttgataaaatgcttcaataatattgaaaaaaggaaga 3730
Db 1727 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTTGAAAAAGGAAGA 1786
Qy 3731 gtagagattcaaacatttcggtgcgccttattcccttttttgcggcattttgccttc 3790
Db 1787 GTATGAGTATTTCAACATTTTCGGGTGTCGCCCTTATTCCTCTTTTTCGGCGCATTTTGCCTTC 1846
Qy 3791 ctggttttgcacccagaacgcgtggtgaagtaaaagatgcgaagatcagttgggtg 3850
Db 1847 CTGTTTTTGCATCCACGAACCGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 1906
Qy 3851 caccagtggtgttacatgaactggatctcaacagcggtaagatccttgagagttttcgcc 3910
Db 1907 CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGTCTTCAGAGTTTCGCC 1966
Qy 3911 ccgaagaacgltttccaatgatgagcacttttaaaagtctgctgatagtggcgcggtattat 3970
Db 1967 CCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTATTTAT 2026
Qy 3971 cccgtattgacgcgcggcaagacaactcgcgtcgccgatacaactattctcagaatgact 4030
Db 2027 CCGGTATTGTAGCGCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAATGACT 2086
Qy 4031 tgggtgagttactcacagctcacagaaagcatcttacggaaggcatgagatgacagtaagaaat 4090
Db 2087 TGGTTGACTACTACCACTCACAGAAAAGCATCTTACGGATGGCATGACATTAAGAAAT 2146
Qy 4091 tatgcagtgtgcataaccatgagtataacactgcggccaacttacttbtgacaacga 4150
Db 2147 TATGCAGTGTCTCCATAACCATAGATGATAACACTTGCGGCCAACTTACTTCTGACAACGA 2206
Qy 4151 tcgagagaccgaagagactaacccgtttttgcacaacatggggggtatcatgtaactccc 4210
Db 2207 TCGGAGGACCAAGAGGCTTAACCGCTTTTTCACAACATGGGGATCATGTAACTCGCC 2266
Qy 4211 ttga tcgttgggaaacggagctgaatgaagcattacccaaagcagcagcgtgacacacga 4270
Db 2267 TTGATCGTTGGAAACCGGAGCTGAATGAAGCATACCAACAGCAGCGGTGACACCAGA 2326
Qy 4271 tgcctgtagcaatggcaaacgltgcgcgaactatttaactggcgaaactacttactctag 4330
Db 2327 TGCTGTAGCAATGGCAACAACGTTTGCSCAAACTATTAACTGGCGGAACACTTACTCTAG 2386
Qy 4331 ctcccgcaacaattaatagactgtaggagcgagataaagtgtcagagaccacttcgc 4390
Db 2387 CTTCCCGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTTCAGGACCACTTCTGC 2446
Qy 4391 gctcgcccttcggctggctgttattgtctgataaaactctggagccggtgagcgtgggt 4450
|||||

Db 2447 GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGACCCGGTGAGCGTGGGT 2506
Qy 4451 ctcgcggtatcatgtcagcactggggccagatggtaagccctcccgtatcgtagttatct 4510
Db 2507 CTGCGGTATCATTTGACGACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 2566
Qy 4511 acagcagcgggagtcaggaactatggtgaacgaataagacagatcgtcgtgagatggtg 4570
Db 2567 ACACGACGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCCTGAGATAGGTG 2626
Qy 4571 cctcactgatlaagcattggtlaactgtcagaccagatttactatataactttagattg 4630
Db 2627 CCTCACTGATTAAGCATTTGTAACCTGTGAGACCAAGTTTACTCATATATATCTTTAGATTG 2686
Qy 4631 atttaaaacttcatttttaatttaaaagatctagtgtaagatcctttttgataactca 4690
Db 2687 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 2746
Qy 4691 tgacaaaatcccttaacgtaggttttccactgagcgcagaccccgtagaaaaaga 4750
Db 2747 TGACCAAAATCCCTTAACGTGAGTTTCTGTTCCACTGAGCGGTACAGCCCGTAGAAAAGA 2806
Qy 4751 tcaaaaggatctcttgagatcccttttctgcgcgtaactcgtcgttgcacacaaaaa 4810
Db 2807 TCAAAAGGATCTTCTTGAGATCCTTTTCTGCGGCTAATCTGCTGCTTGCAAAACAAAA 2866
Qy 4811 aaccacgctaccagcgggtgtgttgcgggatcaagagactcaacactctttttccga 4870
Db 2867 AACCACCGCTACCAGCGGTGTTGTTGCGGATCAAGAGTACCAACTCTTTTTCGGA 2926
Qy 4871 aggttaactggtctcagcagagcgaataacaaaatactcgtccttctagtgtagcgtagt 4930
Db 2927 AGGTAACTGGGTTTACGAGCGCAGATACCAAAATACTGTCTTCTGTGTAGCGGTAGT 2986
Qy 4931 tagggcaccaacttcaagaactctgttagcacgcctacatacctcctctgtctaaacctgt 4990
Db 2987 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTGTCTTAATCCCTGT 3046
Qy 4991 taceagtggtcgtcgcagtgccagtgagataagtcgttcttaccgggttgagactcaagacgt 5050
Db 3047 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTGGACTCAAGACGAT 3106
Qy 5051 agttaccggataaagcgcagcggctcgggtgaacgggggttcgtgcacacagccagct 5110
Db 3107 AGTTACCGGATAAGCGCAGCGGTGCGGTGAACGGGGGTTCGTGCACACAGCCAGCT 3166
Qy 5111 tggagcgaacgacctcacaccgaactgagatacctacagcgtgagctatgagaaagcga 5170
Db 3167 TGGAGCGAAACGACCTTACACCGAACTGAGTACCTACAGCGTGAGCTATGAGAAAGGCCA 3226
Qy 5171 cgtctcccgaaagggaagcggacaggtatccggtaagcggcagggctcggaacagag 5230
Db 3227 CGCTTCCGGAAGGGAAAGGCGGACAGTATCCGTAAGCGCAGGGTTCGGAACAGGAG 3286
Qy 5231 agcgcagcagggagcttcacgggggaacgctggtatctttatagtcctgctcggtttc 5290
Db 3287 AGCGACGAGGAGGCTTCCAGGGGAAACGCTGTGTATCTTTATAGTCTCTGCGGTTTC 3346
Qy 5291 gccactctgacttgagcgtcgatttttctgctgctcgtcagggggggttcgtgcacagccatgga 5350
Db 3347 GCCACCTCTGACTTCAGCGTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCCTATGGA 3406
Qy 5351 aaaaacgcagcaacgcggcctttttacgggttccctggccttttgcctggttttgcctaca 5410
Db 3407 AAAAGCCAGCAACCGCGCTTTTACGGTTCCTGGCCTTTTGTGCTGGCCTTTTGTCTACA 3466
Qy 5411 tgttcttccctgcttatccctgattctgtgataaacgtaattaccgcttttgagtgag 5470
Db 3467 TGTTCTTCTCCGTATATCCCTGATCTGTGGATTAACCGTATTACCCTTTTGTAGTGAG 3526
Qy 5471 ctgataccgctcgcgcagcgaacgcagcgcagcgcagcgcagtcagtcagcgcgagcgg 5530
Db 3527 CTGATACCGCTCGCCCGACCCGACCGACCGCAGCGAGTCAGTGACGAGGAGCGG 3586

QY 5531 agagcgcccaatacgaacccgcctctcccccgcggttgccgattcattaatgcagct 5590
|||||
Db 3587 AAGAGCGCCAATACGAACCCCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCT 3646
QY 5591 ggcagcagaggtttcccgactgaaagcgggagtgagcgacgcaataatgtagtt 5650
|||||
Db 3647 GGCACGACAGGTTTCCCGACTGAAAGCGGCGAGTAGGCGCAACGCAATTAATGTGAGTT 3706
QY 5651 agctcaacttaggcacccaggctttacactttatgcttccggctgtagttgtgtg 5710
|||||
Db 3707 AGCTCACTATTAGGACCCCGAGCTTTACACTTTAAGTTTCGGGCTCGTATGTTGTGTG 3766
QY 5711 gaattgtgagcgatacaatttcacacaggaaacagctatgaccatgattacgcca 5767
|||||
Db 3767 GAATTGTGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 3823

RESULT 3
UC089964
LOCUS UC089964 4473 bp DNA circular SYN 03-SEP-1997
DEFINITION Cloning vector pG5CAT, complete sequence.
ACCESSION U89964
VERSION U89964.1 GI:2293130
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East
Meadow Circle, Palo Alto, CA 94303-4230, USA
East Meadow Circle, Palo Alto, CA 94303-4230, USA. To place an
order call (415) 424-8222 or (800) 662-2566, extension 1.
International customers, please contact your local distributor. For
technical information, call (415) 424-8222 or (800) 662-2566,
extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. If you suspect there is an
error in this sequence, please contact CLONTECH's Technical Support
Department at (415) 424-8222 or (800) 662-2566, extension 3 or
E-mail TECH@CLONTECH.COM

FEATURES
source
1..4473
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
gene
120..779
/gene="CAT"
CDS
120..779
/gene="CAT"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAB68658.1"
/db_xref="GI:2293131"
/translation="MEKIKITGVTVVDSIQWHRKEHFEAFQSAQCYYNOTVOLDITAF
LTKVKNKHYPAPFIHLARLMAHPEFRAMKMGELVINDSVHPCYTVFHEQTFE
SSLWSEYHDDRFQFLHYISQVACVGENLAYFPKGFIEFMFVSANPWSTFSFLNV
AMNDFFAPVETMGKYYTQGDVKVLMPLAIQVHHA VCDGFHVRMLNELQQYCDENQGG
A"
2337..3197
/gene="ampr"
2337..3197
/gene="ampr"
/function="ampicillin resistance"
/codon_start=1

/transl_table=11
/protein_id="AAB68659.1"
/db_xref="GI:2293132"
/translation="MSIOHFRVALIPFFAFCFLPVFAHPETLVKVKDAEDQLGARVGY
TELDLNSKILSEFRRFPNNSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
KSPVTEKHLTDGNTVRELCSAITMSDNTANLLLTITGGKRELFALHNMGDHYTRL
LRWPELNEAFIPNDKSPVYMATLTGLKLTGELLTLASRQQILDIMKADKVAGPL
DLSALPAGWFATPDKSGAGERGSGRTIAALGPDGKPSRIYVIYTTGQATMDERNQIA
ETGASLIKHW"
BASE COUNT 1183 a 1044 c 1046 g 1200 t
ORIGIN
Query Match 45.7%; Score 2637; DB 12; Length 4473;
Best Local Similarity 100.0%; Pred. No.0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3131 cgaattcaactgcccgcgtcttttacaacgctgactggaagaaacccctggcgttaccacaac 3190
|||||
Db 1735 CGAATTCACCTGGCCGTCGTTTACAACGCTGCTGACTGGGAAACCCCTGGCGTTACCCCAAC 1794
QY 3191 ttaatcgccctgcagcacacatccccctttccgcagctggtgtaataagcagagggccgcga 3250
|||||
Db 1795 TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTATATAGCGAAGAGGCCCGCA 1854
QY 3251 ccgatcgcccttcccaacagttgctgcagcctgaatggcgaaatggcgccctgatgcggtatt 3310
|||||
Db 1855 CGCATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATCGCGTATT 1914
QY 3311 ttctccttaacgactctgctgcgtatttcaacacgcgcatatgggtgcactctcagtaaatct 3370
|||||
Db 1915 TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAAATCT 1974
QY 3371 gctctgatccgcatagttaagccagccgcgcacacccgcgaacacccgcgtgagcgccct 3430
|||||
Db 1975 GCTCTGATGCCGATAGTTTAGCCAGCCCGCACACCCGCCACACCCGCTGAGCCGCCCT 2034
QY 3431 gacgggctgtctgctccgggcatccgcttacagacaagctgtgaccgctctccggagct 3490
Db 2035 GACGGGCTTGTCTGCTCCCGGATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 2094
QY 3491 gcatgtgcagaggttttccacgctcatcccgaaacacgcgcgagacgaaagggcctcgta 3550
|||||
Db 2095 GGATGTGTGAGAGGTTTTCACCGTCTATCCCGAAGCCGCGAGACGAAAGGCCCTCGTGA 2154
QY 3551 tacgcctattttataggttaatgctcatgataatagtttcttagcgtcaggtgcga 3610
Db 2155 TAGCGCTATTTTATAGGTTAATGTCATGATATATAGTTTCTTATAGAGTCAAGTGGCA 2214
QY 3611 cttttcggggaaatgtgcggaacccctatttgtttatttttttcaataacatttcaata 3670
|||||
Db 2215 CTTTTCGGGGAATGTGCGCGGAACCCCTATTGTTTATTATTCTTAATACATTTCAATA 2274
QY 3671 tctatccgctcatgagacaataaccctgataaatgctcaataataattgaaagagaga 3730
Db 2275 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAAATATTGAAAAGGAGA 2334
QY 3731 gtagagatttcaacatttccgtgtgcgccttatcccttttttggcgcattttgccttc 3790
Db 2335 GTATGAGTATTCAACATTTCCGCTGTCGCCCTATTTCCTTTTGGGCAATTTGGCTTC 2394
QY 3791 ctgtttttctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3850
Db 2395 CTGTTTTTCTCACCACAGAACGCTGCTGTAAGTAAAGATGCTGAAGATCATGTTGGTG 2454
QY 3851 cgcagatgggttacatcgaactggaatctcaacagcgttaagatccttgaggttttcgcc 3910
Db 2455 CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCTCTGAGAGTTTTCGCC 2514
QY 3911 ccgaagaacgcttttccaatgatgagcacttttaaaagttctgctatgtggcggtattat 3970
Db 2515 CCGAAGAACGTTTTTCCCAATGATGAGCACATTTTAAAGTTCTGCTATCTGCGCGGTATTAT 2574

Qy 3971 cccgtattgacgcggcggaagcaactcggctcgccgcgcatacactattctcagaatgact 4030
|||||
Db 2575 CCCGATTTAGCCCGCGGCAAGAGCAACTCGGTCGCCGCATACATATTCTCAGAAATGACT 2634
|||||
Qy 4031 tgggttagtactaccagtcacagaaagcatcttaggtagtgccatgacagtaagagaat 4090
|||||
Db 2635 TGGTTGAGTACTACCAGTTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAT 2694
|||||
Qy 4091 tatcagtgctgcataaacatagtgatataacactgaggccaaacttactcttgacaaga 4150
|||||
Db 2695 TATCAGTGTCTGACCAATCAACATGATGATAACACTGGCGCCAACTTACTTGTGACAAGA 2754
|||||
Qy 4151 tcggagacgcgaagagtaaccgctttttgacaaacatgaggagcatatgaactcgcc 4210
|||||
Db 2755 TCGAGAGACCGAAGAGCTAACCCGCTTTTGTGCAACATGCGGGATCATGTAACCTGCC 2814
|||||
Qy 4211 ttgatctgtggaaaccggagctgaatgaagccataccaaacagcagcgtgacaccacga 4270
|||||
Db 2815 TTGATCGTTGGGAACCGAGCTGAATGAAGCCATACCAACAGCAGGCTGACACACGA 2874
|||||
Qy 4271 tgcctgtagaataggcaaacagctgacgaactattaaactggcgaactacttactctag 4330
|||||
Db 2875 TGCTGTAGCAATGGCAACAACGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAG 2934
|||||
Qy 4331 ctcccggaacaataatagactggatggagcgagataaagtgtcagagaccactctgc 4390
|||||
Db 2935 CTTCGGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGAGCACCTTCGCG 2994
|||||
Qy 4391 gctcgccctccggctggctgtattgctgataaactggagccggtgagcgtgggt 4450
|||||
Db 2995 GCTCGGCCCTTCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCGGCTGAGCGTGGGT 3054
|||||
Qy 4451 ctgcggtatcatctgagcaactggggcagatggtgaagccctccgctatcgtagttact 4510
|||||
Db 3055 CTCGCGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCTAGTTATCT 3114
|||||
Qy 4511 acagcaggggagtcaggcaactatgatgaacgaataagacagatcgctagataggtg 4570
|||||
Db 3115 ACACAGCGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCCTCGAGATAGGTG 3174
|||||
Qy 4571 cctcactgattgaacattggaactgtcagaccgaagtttactatatactattagattg 4630
|||||
Db 3175 CTTCACTGATTGAAGCATTTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTG 3234
|||||
Qy 4631 atttaaaacttatttaatttaaaaggtactaggtgaagatcctttttgataaactca 4690
|||||
Db 3235 ATTTAAACTTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCA 3294
|||||
Qy 4691 tgacaaaaatcccttaacgtgagtttcgttccactgaagctcagaccgtagaaaaa 4750
|||||
Db 3295 TGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCTACTGAGCGTCAGACCCCTAGAAAAA 3354
|||||
Qy 4751 tcaagagatctcttgagatcctttttctgcgcgtaactgctgcttgcaaaacaaaa 4810
|||||
Db 3355 TCAAGGATCTCTTGAGATCCTTTTCTGCGGTAATCTGCTGTGCAACAAAAA 3414
|||||
Qy 4811 aaccacgctaccagcggtgtgttggtcgggatcaagagctaccaactcttttcoga 4870
|||||
Db 3415 AACCAACGCTACCAAGCGGTGTTTGTTCGGGATCAAGAGCTACCAACTCTTTTCOGA 3474
|||||
Qy 4871 agttaactggcttcagcagcagatatacaataactgtcttctagtgagcgtagt 4930
|||||
Db 3475 AGTTAACTTGGCTTCAGAGAGCGGAGATACCAATATCTCTTCTAGTGTAGCGGTAGT 3534
|||||
Qy 4931 tagggccaccttcaagaactctgtagcagcgctacataactcgctgtgtaactcctgt 4990
|||||
Db 3535 TAGGCCACCACTTCAAGAACTCTGTAGACCGCCTACATACCTCGCTGTGTAATCTCTGT 3594
|||||
Qy 4991 taccagtgctgcctgcagtgccagataagtgctgtcttaccggggttggaactcaagaagat 5050
|||||
Db 3595 TACCAGTGGCTGCTGCCAGTGGCATAACTGCTGCTTACCGGGTTGGACTCAAGACCAT 3654
|||||
Qy 5051 agttaccggataaagcgcagcgggtcggggtgaacggggggttcgtgacacagcccgact 5110
|||||

Db 3655 AGTTACCGGATTAAGCGCAGCGGTCGGCTGAACGGGGGGTTCGTGCAACAGCCAGCT 3714
|||||
Qy 5111 tggagcgaacacactacacccgaactgagatacctacacagcgtgagctatgaaaaagcgcca 5170
|||||
Db 3715 TGGAGCAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 3774
|||||
Qy 5171 cgeitcccgaaaggagaaaggcgcagacaggtatccggttaagcggcagggtcggaacagag 5230
|||||
Db 3775 CGCTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTTCGGAACAGGAG 3834
|||||
Qy 5231 agcgcagagagagcttccaggggaaacgctcgtatcttctatagtcctgcgggttcc 5290
|||||
Db 3835 AGCGACAGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTCTGTCGGGTTTC 3894
|||||
Qy 5291 gccacctctgacttgagcgtcgatttttctgctcaggtcagggggggcggagcctatgga 5350
|||||
Db 3895 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTACGGGGCGGAGCCTATGGA 3954
|||||
Qy 5351 aaaaacgcagcaacgcggccttttttaacggttctcgtgcttctgctgcttttgcacaa 5410
|||||
Db 3955 AAAACGCCAGCAACCGCGCCTTTTACGGTTCTTGGCCTTTTGTGGCTTTTGTCTACA 4014
|||||
Qy 5411 tgtcttctcctgcgttatccctgattctctggtatgataaacgattaccgcttctgagtag 5470
|||||
Db 4015 TGTCTTTCTTGGTATTCCCTGATTCTGTGGATAACCGTATTACCGCTTTTGTAGTGAG 4074
|||||
Qy 5471 ctgataccgctcgcgcgcagcgcgaacgcagcgcagcgcagcgcagcgcagcgcagcgcg 5530
|||||
Db 4075 CTGATACCGCTCGCGCAGCGCAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAAGCGG 4134
|||||
Qy 5531 aagagcgcccaataacgcgaacgcgcctctcccgcgcttgcgcgattcatttaagtgcagct 5590
|||||
Db 4135 AAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATCAGCT 4194
|||||
Qy 5591 ggcacagagcttcccgactggaagcgcgcagcgcagcgcgaacgaacttaagtgcagtt 5650
|||||
Db 4195 GGCAGCAGAGTTTCCCGACTGGAAAGCGGCAGCTGAGCGCAACCAATTAATGTAGTT 4254
|||||
Qy 5651 agtcaactcattaggcaccgcagcgtttacactttctcctcgctcgtatgtgtgtg 5710
|||||
Db 4255 AGCTCACTCATTAGSCACCCAGGCTTTACACTTTTATGCTTCCGGCTCGTATGTTGTG 4314
|||||
Qy 5711 gaattgtgagcggataaacaatttcacagagaacagcgtatgacctattacgcga 5767
|||||
Db 4315 GAATTGTGAGCGGATAACAATTTTACAGAGAAACAGCTATGACCATGATTACGCCA 4371
|||||

RESULT 4

AX191663 LOCUS AX191663 4960 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 29 from Patent WO0149832.
ACCESSION AX191663
VERSION AX191663.1 GI:15209844

KEYWORDS synthetic construct.

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 4960)

AUTHORS Schwenk, F.

TITLE Transduction of recombinases for inducible gene targeting

JOURNAL Patent: WO 0149832-A 29 12-JUL-2001;

ARTEMIS Pharmaceuticals GmbH (DE)

FEATURES Location/Qualifiers

source

1. 4960

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="vector pCMV-I-Cre-pA"

BASE COUNT 1225 a 1213 c 1296 g 1226 t

ORIGIN

Query Match

45.7%; Score 2637; DB 6; Length 4960;

Best Local Similarity 100.0%; Pred. No. 0;				
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	3131	cgaattcaactggcgcgtcgcttttacaacgctgtagactgggaacacccctgctgttaccacaac	3190	
DB	2289	CGAATTCACGTGGCCGTCGTTTACACAGTCGTGACTGGGAAAACCCCTGGCGTTACCCCAAC	2348	
QY	3191	ttaatgccttgacagcaatccccctttccagcgtgcygtaataagcgaagggcccgca	3250	
DB	2349	TTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA	2408	
QY	3251	cgatgcgaccttcccaacagttgcgcagcctgaatgagcaatggcgctgatgcggtatt	3310	
DB	2409	CGATCGCCCTTCCCAACAGTTGCGCAGGCTGAATGGCGAATGGCGCCCTGATGCGGTATT	2468	
QY	3311	tctccttacgcatctgctggttatttccacacgcataTggtgactctcagTacaatct	3370	
DB	2469	TTCCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCT	2528	
QY	3371	gctctgatccgcatagtttaagccagcccccgcacacccgcgaacccgctgacgcgccct	3430	
DB	2529	GCTCTGATGCCGCATAGTTTAAGCCAGCCCGCACACCCGCGCAACACCCGCTGACGCGCCCT	2588	
QY	3431	gacgggctgtctgtcccgcatcgctcttacagacaagctgtgacgctctccggagct	3490	
DB	2589	GACGGGCTTGTCTGCTCCCGGATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT	2648	
QY	3491	gcatgtgcagagggtttccacgcctcatcccgaaacgcgcgagacgaagggccctgta	3550	
DB	2649	GCATGTGTCAGAGGTTTCCACCGTCATACCGAAAGCCGCGAGACGAAGGCCCTCGTCA	2708	
QY	3551	tacgcctattttataggttaatgtcatgataataatggtttcttagacgtcaggtggca	3610	
DB	2709	TACGCCATTTTATAGGTTAATGTATCATATAATAATGGTTTCTTAGACGTCAGGTGGCA	2768	
QY	3611	cttttcggggaaatgctgcggaacccctattgtttatttttcttaaatattcaattcaata	3670	
DB	2769	CTTTTCGGGGAATGTGCGCGGAACCCCTATTGTGTTTATTTTCTAAATACATTTCAATA	2828	
QY	3671	tgtatccgctcatgagacaataaccctgataaagcttcaataataattgaaaaagaga	3730	
DB	2829	TGTATCCGCTCATGAGACAATAAACCTGATAAATGCTTCAATAATATTCAAAAAGGAAGA	2888	
QY	3731	gtagagattcaacatttccgtgtgcgccttattcccttttttggggatatttgccttc	3790	
DB	2889	GTATGAGTATTCAACATTTCCGTGTGCGCCCTATTTCGCCCTTTTTCGGCATTTTGCCTTC	2948	
QY	3791	ctgttttctcaccagaaacgctggtgaagtaaaagatgctgaagatcagttgggtg	3850	
DB	2949	CTGTGTTTCTCACCCAGAAACGCTGGTCAAGTAAAGATGCTGAAGATCAGTTGGGTG	3908	
QY	3851	cacgagtggtttacatcgaaactggatcttcaacagcgggtaagatcccttgagagttttcgcc	3910	
DB	3009	CACGAGTGGGTTACATCGAATCGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC	3968	
QY	3911	ccgaagaacgttttccaatgtagacacttttaagttctgtctatgtgcgcgggtattat	3970	
DB	3069	CCGAAGAAGCTTTCCAAATGATGAGCACATTTTAAAGTTCTGCTATGTGCGCGGTATTAT	3128	
QY	3971	cccgtattgaccccggaagagcaactcgctgcgcgcatataactattctcagaatgact	4030	
DB	3129	CCCGTATTGACCCCGGGGAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACT	3188	
QY	4031	tggttgagtactcaccagtcacagaaaaactcttacggatggcatgacagtaagagaat	4090	
DB	3189	TGTTTGAGTACTCACACAGTCACAGAAAACCATCTTACGGATGSCATGACAGTAAGAGAT	3248	
QY	4091	tatgcagtgctgccaataacccaTgagtgaTaaacactgcggccaaacttactctcgacaacga	4150	
DB	3249	TATGCAGTCTGCATATAACATGAGTGAATACACTGCGGCCCACTTACTTCTGACAACGA	3308	
QY	4151	tcggaggaccgaagagctaaacgcttttttgcaacaactgggggctatgtaactcgcc	4210	

DB	3309	TCGGAGGACCCGAAGGAGCTAACCGCTTTTTCGCACAACATGGGGGATCATGTAACTCGCC	3368	
QY	4211	ttgatctgtgggaaccggagctgaaTgaagccataTaccaaacgacgagcgtgacaccacga	4270	
DB	3369	TTGATCGTTGGGAACCCGGAGCTGAATGAAGCCATACCAACGACGAGCTGACACCACGA	3428	
QY	4271	tgctgtagaatTgcaacaacgctTgcgcaaacTataaactTggaactTaaactTggcaactTactctctg	4330	
DB	3429	TGCTGTAGCAATGGCAACAGCTTTCGCAAACTATTAATCTGGCAACTACTTACTCTAG	3488	
QY	4331	cttcccgccaaacaattaaTtagactgagTgagcgTgataaaagTtcgagaccacttctgc	4390	
DB	3489	CTTCCCGGCAACAATTAATAGACTGATGAGCGCGATAAAGTTTGAGACCACTTCTGC	3548	
QY	4391	gctcgcccttccggctggctggtttattTgctgataaaactTggaagccggtgagcgTgggt	4450	
DB	3549	GCTCGGCCCTTCGGGCTGGCTGGTATTGCTGATAAATCTGGAGCCGCTGAGCGTGGGT	3608	
QY	4451	ctcgcggtatcttgcagcactggggccagatgTgaagccctcccgtatctcgtagttatct	4510	
DB	3609	CTCGCGTATCATTTGACGACTGGGCCAGATGGTAAGCCCTCCCGTATCTGAGTTATCT	3668	
QY	4511	acacgacgggagtcagTcaactatgTgaacgaataTgacagatcgcTgagataggtg	4570	
DB	3669	ACAGAGCGGGAGTCAAGCAACTATGATGAACGAATAAGACAGATCGCTGAGATAGGTG	3728	
QY	4571	cctcactgattaaacattgTtaactTgcagaccaagTttactcaataTacttttagattg	4630	
DB	3729	CCTCACTGATTAAAGCATTGGTAACCTGCAGACCAAGTTTACTCATATATACTTTAGATTG	3788	
QY	4631	atttaaaacttcaatttttaatttaaagatctagTgaagatccctttttgataaactca	4690	
DB	3789	ATTTAAACTTTCATTTTAAATTAAGGATCTAGTGAAGATCCCTTTTGTGAATCTCA	3848	
QY	4691	tgacaaaactccttaacgtgatttctTccactgagcTgcagaccgcgtagaaga	4750	
DB	3849	TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGGTAGAAAAGA	3908	
QY	4751	tcaaggatctcttgagatccctttttctgcgcgTaaactTgctctgcttgcaaacaaaaa	4810	
DB	3909	TCAAGGATCTCTTGAGATCCTTTTCTGCGCTAATCTGCTGCTTGCMAAACAAAA	3968	
QY	4811	agccacgctacacgagcggtggtttgtTtccggatacgaagctacacactcttttccga	4870	
DB	3969	AACCAACGCTTACAGCGGTGGTTGTTGCGCGATCAAGAGCTACCACACTCTTTTCCGA	4028	
QY	4871	aggtaaactggcttcagcagagcgagataTaccaaaactgTcctcttagtTgtagccgtag	4930	
DB	4029	AGGTAACCTGGCTTCACGACAGCGCAGATACCAAAATACTGCTCTTCTAGTGTAGCCGTAGT	4088	
QY	4931	taggcacacacttcaagaactctgtagcaccgcctacaTactctgctctgctaaactctg	4990	
DB	4089	TAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTGT	4148	
QY	4991	taccagtgctgctgccagtgagataaagtctgTcttaccgggttgggactcaagacgat	5050	
DB	4149	TACCAGTGTCTGCTGACGTGGCGATAAGTGTGTTTACCGGGTTTGGACTCAAGACGAT	4208	
QY	5051	agttaacggataagcgcgagcggtcggtgTgaacggggggttcgTgcacacagccccagct	5110	
DB	4209	AGTTACCGGATTAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCT	4268	
QY	5111	tggagcgaacgacctacaccgaactgagataccTacagcgtgagctatgagaaagcgcca	5170	
DB	4269	TGGAGGGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGGCCA	4328	
QY	5171	cgcctcccaaggggagaaagcgagagTatccggtgaagcgaggggtcggaacagggg	5230	
DB	4329	CGCTTCCCAAGGGAGAAAAGCGGACAGGTATCCCGTAAAGCGGAGGGTCGGAACAGGAG	4388	
QY	5231	agcgcaacgagggagctccagggggaacgcctgTtatctttatagTcctgctgggttcc	5290	
DB	4389	ACGGCACGAGGAGGCTTCCAGGGGGAAACGCCGTGCTGATCTTTATAGTCTCTGCGGTTTC	4448	


```
Qy 5291 gccacctgacttgagcgctcagatcttttttgtagtctcgtcaggggggggagcctatgga 5350
Db 4449 GCCACCTCTGACTTGGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGGAGCCTATGGA 4508
Qy 5351 aaagcgacgaaacgagcgctttttacggctccctggccttttctgcttcttctcaca 5410
Db 4509 AAAGCGCAGCAAGCGCGCTTTTACGGTTCCTTGGCCCTTTTGTGGCCTTTTGTCTACA 4568
Qy 5411 tgttcttctcgcttaccctgattctgttggaataaccgtattaccgccttttgagtgag 5470
Db 4569 TGTCTTCTCTGCTGCTATCCCTGATCTGTGTGATACCGGTATTACCGCCTTTGAGTGAG 4628
Qy 5471 ctgataccgtctgcgcagcgcaacgagcagcagcagcagcagcagcagcagcagcagcag 5530
Db 4629 CTGATACCGCTCGCGCAGCGCAACGACCGAGCGAGCGAGCTAGTGAGCGAGGAAGCGG 4688
Qy 5531 aagagcgcccaatacgcgaacccgctctcccccgcgcttgcccgatcattcaatgcagct 5590
Db 4689 AAGAGCGCCCAATACGCANACCGGCTCTCCCGCGCGTGTGGCCGATTCATTAATGCGAGCT 4748
Qy 5591 ggcacagaggttcccgactggaagcgagcgagcgagcgagcgagcgagcgagcgagcgag 5650
Db 4749 GGCACGACAGTTCGCGACTTGAAGCGGCGAGTGAAGCGCAATTAATGTGAGTT 4808
Qy 5651 agctcaactcattagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 5710
Db 4809 AGCTCACTCAATTAGGCAACCGAGCGTTTACACTTTATGCTTCCGCGCTGATGTGTG 4868
Qy 5711 gaattgtgagcggaatacaatttcacagagaaacagctatgaccatgattacgcca 5767
Db 4869 GAATTGTGAGCGGATACAAATTTCCACAGGAACAGCTATGACCAATGATTACGCCA 4925

RESULT 5
AX114861
LOCUS AX114861 5711 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 21 from Patent WO0129208.
ACCESSION AX114861
VERSION AX114861.1 GI:14031803
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Kuehn, R., von Melchener, H. and Altschmied, J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 21 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE); Frankgen Biotechnologie AG (DE)
FEATURES
source
1. .5711
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector prk65"
BASE COUNT 1329 a 1511 c 1617 g 1254 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3131 cgaattcaactggcgctgcttttacaacgtctgactgggaaacccctggcgttacccaac 3190
Db 3040 CGAATTCACTGGCGCTGCTTTTACAAGCTCGTACTGGGAAACCCCTGGCGTTTACCCAAAC 3099
Qy 3191 ttaatcgcttgagacatccctcttcgccagctggcgtaataagagagagcccgca 3250
Db 3100 TTAATCGCTTGGACACATCCCTCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 3159
Qy 3251 ccgatcgcccttcccaacagttgcgcagcctgaatggcggaatggcgctgatgcggtatt 3310
Db 3160 CCGATCGCCCTTCCCAACAGTTGGCGAGCTGAATGGCAATGGCGCTGATGCGGTATT 3219
```

```
Qy 3311 ttctcttaacgcatctgtgcggtattttcacaccgcatatggtgcactctcagtaaatct 3370
Db 3220 TTCTCCTTACGCATCTGTGCGGTATTTTCACACCGCATATGTTGCACATCTCTCAGTACAATCT 3279
Qy 3371 gcttgatgcgcagatagttaagcagcagcccgacaccccgcaaaccccgctgacgcgcct 3430
Db 3280 GCTCTGATGCGCGCATAGTTAAGCCAGCGCCCGACACCCCGCCAACCCCGCTCAGCGCGCT 3339
Qy 3431 gacgggctgtctgctcccgccatccgcttacagacaagctgtgacccgtctcccgagact 3490
Db 3340 GACGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCCGGAGCT 3399
Qy 3491 gcabgttcagaggtttttcacccgtcatcccgaaacgcgcgagacgaaagggcctcgtga 3550
Db 3400 GCATGTGTACAGAGTTTTCACCGTCATCACCGAAACGCGGAGACGAAAGGCGCTCGTGA 3459
Qy 3551 tacgcctattttataggttaattcatgataataagttttcttagacgtcaggtggca 3610
Db 3460 TACGCCATTTTTATAGGTTAATGTATGATAAATGGTTCTTAGACGTACAGGTGGCA 3519
Qy 3611 ctttcgggggaaatgtgcgcggaacccctattgtttatttttctaaatacatccaata 3670
Db 3520 CTTTTCGGGGAATGTGCGCGAACCCCTATTGTGTATTTTCTAATATACATTCAATA 3579
Qy 3671 tgcataccgctcatgagacaataaacctcgtataaaagtctcaataatattgaaaaaggaaga 3730
Db 3580 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA 3639
Qy 3731 gatagtattcaacatttcogtgcgcccttattcccttttttgcggcatttgccttc 3790
Db 3640 GTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCCTCTTTTTCGCGCATTTTGCCTTC 3699
Qy 3791 ctgttttgcctcacccagaaacgcgtggtgaagtaaaagatgctggaagatcagttgggtg 3850
Db 3700 CTGTTTTGTCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 3759
Qy 3851 cagcagtggtgtacatcogaactggatcacaacagcggtaagatccttgagagtttccgcc 3910
Db 3760 CACGAGTGGGTACATCGAATCGAATCGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC 3819
Qy 3911 ccgaagaacgttttccaatgatgagcacttttaaaagtctctatgtgcgcggtattat 3970
Db 3820 CCGAAGAACGTTTTCCAAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 3879
Qy 3971 cccgtattgacgcgggcaagagcaactcgcgcgcgcacatactattctcagaaatgact 4030
Db 3880 CCCGTATTGACCGCGGCAAGAGCANCTCGCTCGCCGATACACTATTCTCAGAAATGACT 3939
Qy 4031 tggttgactactaccagctcagaaaaagcatcttaaggatggtgcagatgacagagaat 4090
Db 3940 TGGTTGAGTACTACCGAGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAT 3999
Qy 4091 tatcagtgctgcataaccatgagtaacactgcgcgcgcgcacacttactcttgacaaaga 4150
Db 4000 TATGAGTGTGCTGCTCAACCATGAGTATAACACTTGGCGGCAACTTACTCTGACAAGA 4059
Qy 4151 tcggagacccgaaggagctaacccgtttttgcacaacatgggggatcatgtaactcgc 4210
Db 4060 TCGGAGNCCCAAGGAGCTAACCCGTTTTTTCACACACATGCGGGATCATGTAACTCGCC 4119
Qy 4211 ttgatcgttgggaacccggagctgaatgaagccataaccaaacgagcgtgacaccaga 4270
Db 4120 TTGATCGTTGGGACCGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACACGA 4179
Qy 4271 tgcctgtagaataaggcaaacagttgcgcaaacatttaactggcgaaactactactctag 4330
Db 4180 TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 4239
Qy 4331 cttcccgcaacaatttaactagactggatggagcggaataaagttgcagagaccactctgc 4390
Db 4240 CTTCCCGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTTGCAGGACCCTTCTGCG 4299
```

QY 4391 gctcggccctccgctggtgctggttattgctgataataatctcagagccggtgagcgtggct 4450
Db 4300 GCTCGGCCCTCCGCGTGGCTGGTGGTATATGCTGATAAATCTCGAGCCGGTGAGCGTGGGT 4359
QY 4451 ctcgcggtatcattgacgactggggccagatgggtaagccctcccgatcgttagttatct 4510
Db 4360 CTCGGGTATCATGTGAGCACCTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 4419
QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgcgtgagataggtg 4570
Db 4420 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 4479
QY 4571 cctcactgattaaagattggttaactgctagacccaagtttaactcaatatatactttagattg 4630
Db 4480 COTCACTGATTAAAGCATTTGGTAACCTGTACAGACCAAGTTTACTCATATATACATTTAGATTG 4539
QY 4631 atttaaaactctattttaaattaaagatctaggtgaagatccttttggataactctca 4690
Db 4540 ATTTAAACTCTCATTTTANTTTAAAGGATCTAGTGGAAGATCTCTTTTGATAATCTCA 4599
QY 4691 tgacccaaatcccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaaa 4750
Db 4600 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGACGCTCAGACCCCGTAGAAAAA 4659
QY 4751 tcaaggatctctctgagatccttttttctcgcgtaactctgctgctgcaaaaacaaa 4810
Db 4660 TCAAAAGGATCTTTCTGAGATCCTTTTCTGCGGTAACTCTGCTTGTCAAAACAAAA 4719
QY 4811 aaccacgctacacgggtggttctgctcggatcaagagctaccaactcttttccga 4870
Db 4720 AACACCGCTACACGGGTGGTTGTTGCCGGATCAAGAGCTACCACTCTTTTCCGA 4779
QY 4871 agtgaactggtctcagcagcgcagatcaccaataactgtctctcttagtgagccgtagt 4930
Db 4780 AGGTAACCTGGCTTACGACAGCGCAGATACCAAAATACTGTCTCTTAGTGTAGCCGTACT 4839
QY 4931 taggcacacacttaagaactctgtagacccgcctacatacactcgtctgtaactctgt 4990
Db 4840 TAGGCACACACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAACTCCTGT 4899
QY 4991 taccagtgctgctgcagtgccgataagctgctctaccgggttggaactcaagacgat 5050
Db 4900 TACCAGTGCTGCTGCCAGTGGCGATAGCTGCTGTCTTACCGGGTGGACTCAAGACGAT 4959
QY 5051 agttaccggataagcgcagcgtggtggtgaacgggggttcgtgacacagccagct 5110
Db 4960 AGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGCACACAGCCAGCT 5019
QY 5111 tggagcgaacgacctacacgcaactgagatacctacagcgtgagctatgagaaagccca 5170
Db 5020 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 5079
QY 5171 cgtctccgaagaggagaagcgcagagtatccggttaagcgcaggggtcgaaacagag 5230
Db 5080 CGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGGTGGCAACAGGAG 5139
QY 5231 agcgcacgaggagctccagggggaacgcgtggtatctttatagctcgtcggtttc 5290
Db 5140 AGCGCACGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGTCTCTCGGGTTTC 5199
QY 5291 gccactctgactgagcgtcagattttgtgctcgcgcagggggggcgagccctatgga 5350
Db 5200 GCCACTCTGACTTGAGCCTCGATTTTGTGATGTCTGCTGTCGTACGGGGGCGGAGCCCTATGGA 5259
QY 5351 aaaaagcgaacgcggccttttactggtctcctggccttttggccttttggctcctcaca 5410
Db 5260 AANAAGCCAGCAACGGGGCTTTTACGGTTCTGGCCCTTTTGTGTCGCTTTTGTCTACA 5319
QY 5411 tgttcttctcgttataccctgattctctgtgataaacgctattaccgcttttgagtgag 5470
Db 5320 TGTCTCTTCTCGTATATCCCTGATCTCTGTGATTAACCGTATTAACCGCTTTGAGTGAG 5379
QY 5471 ctgataccgctcgcgcagcgaacgacgagcgcagcagtcagtgagcaggaagcgg 5530

Db 5380 CTGATACCGCTCGCGCAGCCGAGCGAGCGAGTCAGTGACGAGGAGCGG 5439
QY 5531 aagagcgcccaatacgcgaacccgctctcccgcgctgttggcgtatcatttaagcagct 5590
Db 5440 AAGAGGCCCATACGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTATGAGCT 5499
QY 5591 ggcagacagggttcccgactgaaagcgggagtcgagcgaacgaatgaatgagtt 5650
Db 5500 GGCAGACAGGTTCCGACTGAAAGCGGAGTGAGCGCAACGCAATTAATGTGAGTT 5559
QY 5651 agctcaactataggcaccccaaggctttcacactttatgtctccggctgtagttgtg 5710
Db 5560 AGCTCACTCATTAGGCACCCAGGCTTTACACTTATGTCTCCGCTGTATGTTGTG 5619
QY 5711 gaattgtgagcggaataacaatttcacaggaacagcgtatgacctgattcgcga 5767
Db 5620 GAATTGTGAGCGGATACAAATTTTCACAGGAACAGCTATGACCATGATTAGGCA 5676

RESULT 6
AX191664
LOCUS AX191664 7332 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 30 from Patent WO0149832.
ACCESSION AX191664
VERSION AX191664.1 GI:15209845
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 7332)
AUTHORS Schwenk, F.
TITLE Transduction of recombinases for inducible gene targeting
JOURNAL Patent: WO 0149832-A 30 12-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
source
1..7332
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pCMV-I-beta-pA"
BASE COUNT 1718 a 1895 c 1964 g 1755 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 7332;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaactgcccgtcgcttttacaacgctgctgactgggaaacccctggcgttaccacaac 3190
Db 4661 CGAATTCACTGGCCGCTCGTGTTTACAACTGCTGCTGTTGGGAAACCCCTGGCGTTACCCCAAC 4720
QY 3191 ttaactgccttcagcacatccctcttcgacgctgctgctgtaataagcgaagggcccgca 3250
Db 4721 TTAATGCGCTTCGAGCACATCCCTCTTCGCGAGCTGGCGTAAATAGCGAAGAGGCCCGCA 4780
QY 3251 ccgactgcctctcccaacagttgcgagcctgaatggcgaatggcgcctgagcgtgatt 3310
Db 4781 CCGATCGCCCTTCCCAACAGTTGGCGAGCTGAATGCGCAATGGCGCTGATCGCGTATT 4840
QY 3311 ttctcttcagcatctgtcgggtatttcacacgcgatatggtgcactctcagtaacaatct 3370
Db 4841 TTCTCCTTACGCATCTGTGCGGTATTTTACACGCGCATATGGTGCATCTCAGTACAACTCT 4900
QY 3371 gctctgatacccatagtttaagccagcccgacaccccgcaacacccgctgagcgccct 3430
Db 4901 GCTCTGATCCGCATAGTTAAGCCAGCCCGACACCCCGCAACACCCGCTGAGCGCCCT 4960
QY 3431 gacgggctgtctgctcccggtcatccgcttacagacaagctgtgaccgtctcctcggagct 3490
Db 4961 GACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 5020
QY 3491 gcatgtgtcagagggtttttcaccgctcatcaccgaaacgcgcgagagcaagggcctcgtga 3550

Db 5021 ||||| GCATGTGTCAGAGGTTTTCACCGTCATCACGGAACGGCGGAGACGAAGGCGCTCGTGA 5080
 Qy 3551 tacgcctattttataggttaagtctatgataataatggtttcttagcgtcagtgga 3610
 Db 5081 TACGCCATATTTTATAGGTTAATGTCATGATAATATGTTTCTTAGACGTCAGGTGCA 5140
 Qy 3611 cttttcggggaaatgctgcgggaaacccctatttggttatttttctaaatacaata 3670
 Db 5141 CTTTTCGGGGAAATGTGCGGGAACCCCTATTTGTTTATTTTCTTAAATACATTCANAATA 5200
 Qy 3671 tqtatccctcatagagacaataaacctctgataaactctcaataatgataaaaggaaga 3730
 Db 5201 TGTATCCGCTCATGAGACATAAACCTGATAAATGCTTCAATAATATTTGAAAAAGGAAGA 5260
 Qy 3731 gtabagattcaacatttcogtgcgcccatttcccttttttgcggcatttttccttc 3790
 Db 5261 GTATGAGTATTCACATTTCCGTGCGCCCTTATTTCCCTTTTTCGGGCATTTTGCCTTC 5320
 Qy 3791 ctggttttgcctcaccagaaacgctgggtgaaagttaaagataaagatcagttgggtg 3850
 Db 5321 CTGTTTTTGTCTACCCAGAAACGCTGGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG 5380
 Qy 3851 cagcagtggtgtacatcgaaactggatctcaacagcggtaagatccttgagagtttcgcc 3910
 Db 5381 CACGAGTGGTTACATCGAATCGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC 5440
 Qy 3911 ccgaagaacgttttccaatgatgagcacttttaagttctgctatgctggtgcggtattat 3970
 Db 5441 CCGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGTATGTGGCGCGGTATTAT 5500
 Qy 3971 cccgtatgagccggggaagagaactcgtgcgcgatacaactattctcagaatgact 4030
 Db 5501 CCGGTATTGTACCGCGGGGAAGAGCAACTCGGTGCGCGGATACACTATTCTCAGATGACT 5560
 Qy 4031 tggttgagtactcacacagtcacagaaaagcatcttcaggtatggcagatgaagagaat 4090
 Db 5561 TGGTTGAGTACTCACAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAT 5620
 Qy 4091 tatgcagtctgcccataaccatgagtatgaacactgcggccaacttactctgacaaga 4150
 Db 5621 TATGCAGTCTGCCATAACCATGAGTGATTAACACTCGCGCCAACCTTACTTCTGACAACGA 5680
 Qy 4151 tcggaggcgaagagctaaacgcttttttgcaacaactgggggatactgtaactcgcc 4210
 Db 5681 TCGGAGGACCGAAGGAGCTAAACCGTTTTTTCACAACATGGGGGATCATGTAACTCGCC 5740
 Qy 4211 ttgatcggtgggaacggagctgaatgaagccataccaaacagcagcgtgacacacga 4270
 Db 5741 TTGATCGTTGGCAACGGGAGCTGAATGAAGCCATACCAAAACGACGCGTGACACACGA 5800
 Qy 4271 tgcctgtagcaatggcaacaacgcttgcgaacactttaactggcgaacttactcttag 4330
 Db 5801 TGCCCTGTAGCAATGGCAACAACGTTTGGCGCCAGATGGTAAGCCCTCCCGTATCTATCT 5860
 Qy 4331 ctctccggcaacaataatagactgagtgaggcggaataaagtctcaggaccaactctgc 4390
 Db 5861 CTTCGGGCAACAAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGAGCCACTTCGCG 5920
 Qy 4391 gctcgcccttcggcgtggtgtttattgctgataaaactgagcgggtgagcgtgggt 4450
 Db 5921 GCTCGGCCCTTCGGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGCTGAGCGTGGGT 5980
 Qy 4451 ctgcgggtatcatatgcagcacttggggccagatggtaagccctcccgatctagttatct 4510
 Db 5981 CTTCGGGATCATATGCAGCACTTGGGCGCCAGATGGTAAGCCCTCCCGTATCTATCT 6040
 Qy 4511 acacagcgggagtcaggcaactatgatatgaagcaaatagacagatcgtctgagatgg 4570
 Db 6041 ACACAGCGGGAGTCAGGCAACTATGGATGAAACGAATAAGACAGATCGCTGAGATGGTG 6100
 Qy 4571 cctcactgattaaagcattggtaactgcagaccaggttactcatatataacttagattg 4630

Db 6101 CCTACTGATTAAGCATTTGGTAACCTGTGAGACCAAGTTTACTCATATATATCTTTAGATTG 6160
 Qy 4631 atttaaaacttcatttttaatttaaaagatctaggtgaagatcccttttttgataatctca 4690
 Db 6161 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 6220
 Qy 4691 tgaccaaaatcccttaacgctgaggttttcctcactgagcgtcagaccccgtagaaaaga 4750
 Db 6221 TGACCAAAATCCCTTAACGTGAGTTTCTGTCACGTGAGCGTCAGACCCCGTAGAAAAGA 6280
 Qy 4751 tcaaaagatctctcttgagatcccttttctcgcgtaactctgctgctgcaaaacaaaa 4810
 Db 6281 TCAAGAGATCTCTTGTGAGATCCTTTTTCGCGCGTAATCTGCTGCTGTCGCAACAAAA 6340
 Qy 4811 aaccaccctaccagcgggtggtttgttgcgggataaagagctacccaactcttttccga 4870
 Db 6341 AACCAACCGCTACAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGA 6400
 Qy 4871 aggttaactggcttcagcagagcgagatacaaaaactctcctcttagtctagcgtagt 4930
 Db 6401 AGGTAACTGGCTTTCAGCAGAGCGCAGATACCAAAATACTGTCCTTCTAGTGTAGCCGTAGT 6460
 Qy 4931 taggcacacacttcaagaactctgtagcacgcctacatacactcgtctctgctlaactctgt 4990
 Db 6461 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTGCTTAATCTGT 6520
 Qy 4991 taccagtggtcgtccagtggtgagataagtcgtcttaccgggttggaactcaaacagat 5050
 Db 6521 TACCAGTGGCTGCTGCCAGTGGCGATAGTCGTGCTTACCGGGTGGACTCAAGACGAT 6580
 Qy 5051 agttaccggataagcgcagcgggtcgggtgaacgggggttcgtgcacacagccagct 5110
 Db 6581 AGTTACCGGATTAAGCGCAGCGGTGGGGTGAACGGGGGGTTCGTGCACACAGCCAGCT 6640
 Qy 5111 tggagcgaacgacctacacccgaactgagatacctacagcgtgagctatagaaagcgcca 5170
 Db 6641 TGGAGGGAACGACCTACACCGAACTGAGATACTACAGCGTCAGCTATGAGAAAGCGCCA 6700
 Qy 5171 cgtctccgaaggggaaagcggacaggtatccggtaagcgcaggggtcggaacaggg 5230
 Db 6701 CGTTCGGGAGGGAAGGCGGACAGGTATCCGGTAAGCGCAGGGTCGGAACAGGAG 6760
 Qy 5231 agcgcagcagggagcttccaggggaaacgctggtatcttattagctcgtcggttttc 5290
 Db 6761 AGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGTCTCTGCGGGTTTC 6820
 Qy 5291 gccaactctgacttgagcgtcgtattttgtgctcgtcaggggggagcggcctatgga 5350
 Db 6821 GCCACTCTGACTTGAGCGTCCATTTTGTGATGCTGCTCAGGGGGCGGAGCCCTATGGA 6880
 Qy 5351 aaaaagcagcaacgcggcctttttacgggttccctggccttttgccttttgcctcaca 5410
 Db 6881 AAAAGCCCAACGACGGCCTTTTACGGTTCTCGGCCCTTTTGTGCGCTTTTGTCTACA 6940
 Qy 5411 tgttttctcgtcttaccctgattctgtgataaacgattaccgcttttgagtgag 5470
 Db 6941 TGTCTTCTCCTCGGTATCCCTGTATCTGTGGATAACCGTATTACCGCTTTTGTAGTGAG 7000
 Qy 5471 ctgataccgctcgcgcagccgaacagcagcagcagcagtcagtgagcggaggaagcgg 5530
 Db 7001 CTGATACCGCTCGCGCAGCCGCAACGACGAGCGACGCGAGCTCAGTGACGGAGGAAGCGG 7060
 Qy 5531 aagagcgcacaatagcgaacacgcctctcccgcgcttggccgatttcaatcagct 5590
 Db 7061 AAGAGCGCCCAATAGCAAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAATGACGCT 7120
 Qy 5591 ggcacagcaggtttcccgactgaaagcggcgagtgagcgaacgcaataatgtgagtt 5650
 Db 7121 GGCACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTT 7180
 Qy 5651 agctcactcattaggcaccacccagcttttaccactttatgcttcggtcgtatgtgtg 5710
 Db 7181 AGCTCACTCATTTAGGCACCCCGAGGCTTTACACTTTATGCTTCGGGCTCGATGTGTGTG 7240

QY 5711 gaattgtgagcgagataacaatttcacacagggaaacagctatgacccatgattacgcga 5767
|||||
Db 7241 GAATTGTGAGCGGATAACAATTTACACAGAGAAACGCTATGACCATGATTAGGCCA 7297
|||||

RESULT 7
AX114872/c
LOCUS
DEFINITION
AX114872
ACCESSION
AX114872
VERSION
AX114872.1
KEYWORDS
GT:14031814
synthetic construct.
SOURCE
synthetic construct.
ORGANISM
artificial sequence.
REFERENCE
1 (bases 1 to 8062)
AUTHORS
Kuehn, R., von Melchener, H. and Altschmied, J.
TITLE
Conditional gene trapping construct for the disruption of genes
JOURNAL
Patent: WO 0129208-A 32 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE); Frankgen Biotechnologie AG (DE)
FEATURES
Location/Qualifiers
1..8062
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector PRK76"
BASE COUNT 1960 a 2036 c 2121 g 1945 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 8062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaactgcgctgttttacaacgctgcgtactggagaaacccctggcgttacccaac 3190
|||||
Db 8045 CGAATTCACTGGCGCTGCTTTTACAACGCTGCTGACTGGGAAAACCCCTGGCGTTACCCAAC 7986
|||||

QY 3191 ttaatgccttgcagcacatcccccttttcgacagctggcgtaatagcgaagagcccgca 3250
|||||
Db 7985 TTAATCGCCTTGACGACATCCGCCCTTTGCGCAGCTGGCGTAATAGCAAGAGGCCGCA 7926
|||||

QY 3251 ccgatacgccttcccaacagctgcgagcctgaatggcgaaatggccctgagcgtatt 3310
|||||
Db 7925 CGATCGCCTTCCCAACAGTTGCGCAGCCTGAATGGCAATGGCGCCTGATCGGTATT 7866
|||||

QY 3311 ttctcttacgcactgtgcggtatttcacacgcgcatatggtgcactctcagtaaatct 3370
|||||
Db 7865 TTCTCCTTAGCATCTGTGGGTATTTCACACCGCATATGGTGCACTCTCAGTACATCT 7806
|||||

QY 3371 gctctgatgccgcatagttaagccagccccgcacacccgcacacccgctgacgcgccct 3430
|||||
Db 7805 GCTCTGATGCCGATAGTTAAGCCAGCCCGCACACCCCGCCACACCCGCTGACGCCCT 7746
|||||

QY 3431 gacggcctgtctgtcccggaatccgcttacagacaagctgtgacccgtctcggagct 3490
|||||
Db 7745 GACGGCCTGTCTGTCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGAGCT 7686
|||||

QY 3491 gcatgtgtcagaggttttaccgtctatcccgaaacgcgcgagacgaagggcctcgtga 3550
|||||
Db 7685 GCATGTGTACAGAGTTTTCACCGTCATACCGNAACCCGCGAGACGAAGGGCCTCGTGA 7626
|||||

QY 3551 tacgcctattttatagggttaatgcatgataataatgggtttcttagacgtcaggtgca 3610
|||||
Db 7625 TAGCCCTATTTTATAGTTAATGTCTATATAATATGTTTCTTAGACGCTCAGGTGGCA 7566
|||||

QY 3611 cttttcggggaaatgscgggaacccctattgttttatttttttcttaaaacattcaataa 3670
|||||
Db 7565 CTTTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTATTTTCTTAATATACATTCAATA 7506
|||||

QY 3671 tgtatccgctcatgagacataaccctgataaatgcttcaataaatatgaaaaaggaga 3730
|||||
Db 7505 TGATCCGCTCATGACACAATAACCCGTATAATGCTTCAATTAATATTGAAAAGGAAGA 7446
|||||

QY 3731 gtatgagtattcaacattttccgtgtgcgccttattcccttttttggcgcaattttgcttc 3790
|||||
Db 7445 GTATGAGTATTCAACATTTCCGTGTGCGCCTTATTCCCTTTTTCGGGCATTTTGCCTTC 7386
|||||

QY 3791 ctgtttttgctcaaccagaaacgctggtgaaagtaaaagatgcgtgaagatcagttgggtg 3850
|||||
Db 7385 CTGTTTTTGTCTACCCAGAAACGCTGTTGAAAGTAAAGATGCTGAAGATACAGTTGGGTG 7326
|||||

QY 3851 cacgagtgggtttacatcgaactgagatctcaacagcggttaagatcccttgagagtttccgc 3910
|||||
Db 7325 CACGAGTGGTTACATCGAACTCGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC 7266
|||||

QY 3911 ccgaagaacggttttccaatgatgagcaccttttaagttctctgctatgtggcgcggtattat 3970
|||||
Db 7265 CCGAAGAACGTTTTTCCAATGATGAGCACTTTTTAAGTTCTGCTATGTGGCGCGTATTAT 7206
|||||

QY 3971 ccgattatgacgcggcggaagagacaactcgtgcgcgcatacaactattctcagaatgact 4030
|||||
Db 7205 CCCGATTGACGCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACT 7146
|||||

QY 4031 tggttgagtactcaccagtcacagaaagcatctctacgagatgcacagtaagagaat 4090
|||||
Db 7145 TGGTTGAGTACTCACCAGTCACAGAAAAGCTCTTACGGATGCGATGACAGTAAGAGAAT 7086
|||||

QY 4091 tatgcagtgctgcataaaccatgagtataaacactcggcgcaacttactcttgacaacga 4150
|||||
Db 7085 TATGCAGTGTGTCATTAACCATGAGTGATAACACTCGGGCCAATTACTTCTGACAAGGA 7026
|||||

QY 4151 tcggagagcgaagagagctaacccgcttttttgcaaacatcgggggatactgtaactcgc 4210
|||||
Db 7025 TCGGAGAGCCGAAGGAGCTTAACCGCTTTTGTGACAACTATTAACTGGCGGATCATGTAACTCGCC 6966
|||||

QY 4211 ttgactgttgggaaccgagctgaatgaagccataccaaacgacgagcgtgacaccacga 4270
|||||
Db 6965 TTGATCGTTGGGAACCGGAGCTCAATGAAGCCATACCAACGACGAGCGTGACACCACGA 6906
|||||

QY 4271 tgctgtgacaaatggcaacaacgctgcgcaactataaactggcgcaacttacttactctag 4330
|||||
Db 6905 TGCTGTAGCAATGGCAACAACCTTGGCGAATCTATTAACTGGCGAATCTACTTACTCTAG 6846
|||||

QY 4331 cttcccgcaacaattaatagactgagtgagcggaataaagtgcaggagcaactctcgc 4390
|||||
Db 6845 CTTTCCCGGCAACAATTAATAGACTGGAGCGCGGATAAAGTTGAGGACCACTTCTGTC 6786
|||||

QY 4391 gctcgcccttccgctggcgtgttttattgctgataaaactggagccggtgagcgtgggt 4450
|||||
Db 6785 GCTCGGCCCTTCGCGCTGGCTGGTTTATTGCTGATAAATCGGAGCGGTGACCGTGGGT 6726
|||||

QY 4451 ctcgcggtatcatgacgcaactggcgccagatggtaagccctcccgctatgtagttatct 4510
|||||
Db 6725 CTGCGGTATCATTTGACGACCTGGGCGCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6666
|||||

QY 4511 acaacgaggggagtcaggaacataatgataaagaataagacagatcgctgagataagggtg 4570
|||||
Db 6665 ACACGAGGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGGTGAGTAGGTG 6606
|||||

QY 4571 cctcactgatgaagcattggttaactgtcagaccaagtttactcatataactttagatg 4630
|||||
Db 6605 CCTCACTGATTAGCATTTGGTAACCTGTACAGCCCAAGTTTACTCATATATATTTAGATTG 6546
|||||

QY 4631 atttaaaacttcatttttaatttaaagatctagtggaagatccctttttgataatctca 4690
|||||
Db 6545 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAAATCTCA 6486
|||||

QY 4691 tgacaaaaatcccttaacgtgagtttctccactgagcgtcagagcccgtagaagaaga 4750
|||||
Db 6485 TGACAAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGGTAGAAAAGA 6426
|||||

QY 4751 tcaagagctctcttgagatccctttttctgcgcttaactctgctctgtcgtcaaacaaaaa 4810
|||||
Db 6425 TCNAAGGATCTCTTCAGATCCTTTTTCGCGCGCTAATCTGCTGCTTGCACAAACAAAAA 6366
|||||

```
QY 4811 aaccacgcctaccagcggtggtttgttgcgagatcaagagctaccacactctttttccga 4870
Db 6365 AACACCCGCTACCAAGCGGTGGTGTGTCGGGATCAAGAGCTACCAACACTCTTTTCCGA 6306
QY 4871 agttaactggtcttcagcagcgcagatcaccaataactgtctctctagttagcgtagt 4930
Db 6305 AGGTAACGTGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGAGCGGTAGT 6246
QY 4931 taggcacccacttaagaactctgtagacacgcctacactcctcgtctgtaactcgt 4990
Db 6245 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACTCGCTCTGTAATCTCTGT 6186
QY 4991 taccagtggtcgtcgcagtggcagataagtcgtgtctctaccgggttgactcaagacgat 5050
Db 6185 TACCAGTGGCTGCTGCCAGCGGCGCATGAAGTCGTCTTACCGGGTTGGACTCAAGACGAT 6126
QY 5051 agttaccggataaggcgacggtgcgggctgaacggggggttcgtgcacacagcccgact 5110
Db 6125 AGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCT 6066
QY 5111 tggagcgaacacacacacgaactcagatcacctacacagcgtgaactatgagaaagcgcca 5170
Db 6065 TGGAGCAACGACACTACCCGCACTAGATACCTACAGCGGTGAGCTATGAGAAAGCGCCA 6006
QY 5171 cgtctccgaagggagaaagcgacaggtatccggtaaagcgagcaggtcggaacacagag 5230
Db 6005 CGCTTCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGCTGGAACAGAGAG 5946
QY 5231 agcgacaggggagcttccaggggggaaacgcctggtatctttaagtcctgctgggttc 5290
Db 5945 AGCGCAGAGGAGCTTCCAGGGGGAAACGCTGATCTTTATAGTCCCTGTCGGGTTTC 5886
QY 5291 gcaacctgaactgagcgtogaattttgtgatcgtcgtcagggggcgagcctatgga 5350
Db 5885 GCCACCTCTGACTTGAAGCGTGAATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGGA 5826
QY 5351 aaaaacgacgaacgcggtcttttacggttcctgagccttttgccttttgcctcaaca 5410
Db 5825 AAAACGCCAGCAACCGCGCTTTTACGGTTCCTGGCCCTTTTGTGCGCTTTTGTCTCACA 5766
QY 5411 tgttcttctcgtgtatccccgtattctgtggataaccgctattaccgcctttgagtgag 5470
Db 5765 TGTTCTTCTCGCTTATCCCTGATCTGTGGATAACCGTATTAACCGCTTTTGTGAGTGAG 5706
QY 5471 ctgatacgcgtcgcgcagcgcgaacgcgcagcgcagcgcagcgcagcgcagcgcagc 5530
Db 5705 CTGATACCGCTTCGCGCGCAGCGGAACCGGACCGAGCGCAGCTCAGTGAGCGAGGAACGG 5646
QY 5531 aagagcgcccaatacgaacccgctctccccgcgcttgccgagcattcattaatgcagct 5590
Db 5645 AAGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTGGCCGATTCATTAATGCGAGCT 5586
QY 5591 ggacacagaggtttcccgactggaagcgggagtgagcgcgaacgaactaatgtgagtt 5650
Db 5585 GGCACGACAGGTTTCCGACTGGAAAGCGGCGACTGAGCGCAACGCAATTAATGTGAGTT 5526
QY 5651 agctcaactataggaccaccagcgtttacactttatgcttcggcgctgtagttgtgtg 5710
Db 5525 AGCTCACTCATTAGGACACCCCGAGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTG 5466
QY 5711 gaattgagcggataaacaatttcacacaggaacacagctatgacacattgatacga 5767
Db 5465 GAATTGTAGCGGATAACAATTTTACACAGGAAACAGCTATGACCATGATTACGCCA 5409

RESULT 8
AX114871/c
LOCUS AX114871 8153 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 31 from Patent WO0129208.
ACCESSION AX114871
VERSION AX114871.1 GI:14031813
KEYWORDS
SOURCE synthetic construct.
```

```
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 8153)
AUTHORS Kuehn, R., von Melchner, H. and Altschmied, J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 31 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE); Frankgen Biotechnologie AG (DE)
FEATURES
    source
        location/Qualifiers
            1..8153
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="vector prk74"
BASE COUNT 1976 a 2133 c 2067 g 1977 t
ORIGIN
Query Match 45.7%; Score 2637; DB 6; Length 8153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3131 cgaattcacctggcgtctgttttacacgctgctgactgggaaacccctggcgttaccacca 3190
Db 8136 CGAATTCACTGGCGCTGCTTTTACACGCTCTGACTGGGAAACCCCTGGCCTTACCCCAAC 8077
QY 3191 ttaatgccttgacagacatcccccttttgcacagctggcgttaataagagaagaccgcga 3250
Db 8076 TTAATGCCTTGCAGACATCCCTTTTCGACGCTGGCGTAATAGGAAAGAGGCCCGCA 8017
QY 3251 ccgatcccccttcccaacagttgcgagcctgaatgcgaatgcgctgagcgtgatt 3310
Db 8016 CCGATCGCCCTTCCACAGCTTGCAGCCCTGAATGGGAATGGCGCTGATGGGTATT 7957
QY 3311 ttctccttagcactctgtgcggtatttcacaccgcataatggtgcactctcagtaacact 3370
Db 7956 TTCTCCTTACGCATCTGTGCGGTATTTTACACACCGCATATGTTGCACCTCAATACT 7897
QY 3371 gctctgtagcgcagtagttaagccagccccgcacccccgcacccccgcgcgcgcct 3430
Db 7896 GCTCTGATGCGGCATAGTTTAAAGCCAGCCCGCACACCCGCCAACCCGCTGACGCGCCT 7837
QY 3431 gacggggtgtcgtctcccgccatccgcttacagacacagctgtgacgctcccgaggct 3490
Db 7836 GACGGGCTTGTCTGCTCCCGCATCCGCTTACACACAGCTGTGACCGCTCTCCGGAGCT 7777
QY 3491 gcatgtgcagaggttttccacgcgtcaaccgaaacgcgcgagacgaagggcctcgtga 3550
Db 7776 GCATGTGTCAGAGGTTTTCACCGTCATCACGAAACGCGCAGACGAAAGGCGCTCGTGA 7717
QY 3551 tacgcctattttataggttaagtcaataaagggtttcttagacgtcaggtggca 3610
Db 7716 TACGCCCTATTTTATAGGTTAATGTCATGATTAATGGTTTCTTAGACGTCAGGTGGCA 7657
QY 3611 ctttcggggaagtgcgcgggaacccctattgtttatttttcttaatacatcaata 3670
Db 7656 CTTTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTTCTTAATACATTCAATA 7597
QY 3671 tgtatccgctcatgagacaataaacccctgataaatgcttcaataatattgaaaaaggaaga 3730
Db 7596 TGTATCCGCTCATCAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA 7537
QY 3731 gtatgatttcaacattccgctgcgccttatcccttttttgcggcatttgccttc 3790
Db 7536 GTATGAGTATTCAACATTTTCGCTGCGCCCTTATTCCTTTTTCGGGCATTTTGCCTTC 7477
QY 3791 ctgttttgcctccacagaaacgctggtgaagtaaaagatgctgaagatcagttgggtg 3850
Db 7476 CTGTTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 7417
QY 3851 cagcagtggtttacatcgaaactggtatctcaacacgggtcaagatccttgagagttttcgcc 3910
Db 7416 CACGAGTGGGTACATCGAACTGGATCTCAACAGGGGTAAAGATCCTTTCGAGAGTTTCGCC 7357
QY 3911 ccgaagaacgcttttccaatgatgagcacttttaagttctctgtatgtggtcggtlata 3970
```


REMARK revised by [3]
REFERENCE 3 (bases 1 to 2746)
AUTHORS Technical, Services.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2000) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT On Jan 27, 2000 this sequence version replaced gi:58207.
See X5300-X5335 for related vector sequences
This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.

FEATURES
source location/Qualifiers
1. .2746
/organism="Cloning vector pGEM-42"
/db_xref="taxon:90111"
misc_feature 1
note="SP6 transcription initiation site"
7. .63
note="multiple cloning sites"
58. .87
note="T7 promoter"
70
misc_feature 1
note="T7 transcription initiation site"
96. .325
note="lac operon sequence"
complement(106..122)
primer_bind
note="pUC/M13 reverse sequencing primer"
complement(108..110)
gene
note="lacZ"
complement(108..110)
misc_feature 1
note="lacZ"
130..146
note="lacZ start codon"
misc_feature 1
note="lac operator"
complement(1267..2127)
gene
note="bla"
complement(1267..2127)
CDS
note="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="CAA46397.2"
/db_xref="GI:6782315"
/transl_table="MSIQHFRVALIPFAAFCLPVPFAHPETLVKVKDAEDQLGARVGY
IELDNSKILESFRPEHPFMSTFKVLLGAVLSRIDAGQEOELGRRIHYSNDLVE
YSPVTEKHLTDGTVRELCNAATMSDNTAANLLTTIGGPKELTAFLHNGDHVTRL
DRWPELNEAIPNDERITTPVAMATTLRLKLLTGELLTLASRQQLIDWMDKRVAGPL
LRSLALPAGWFIADKSGAGENGSKGIIAALGPDGKPSRIVIVITGTSQATMDERNRQIA
EIGASLIKHW"
2563..2726
misc_feature 1
note="lac operon sequence"
2686..2702
primer_bind
note="pUC/M13 forward sequencing primer"
join(2730..2746;1..3)
promoter
note="SP6 promoter"
684 a 684 c 699 g 679 t
BASE COUNT
ORIGIN

Query Match 45.78; Score 2636; DB 12; Length 2746;
Best Local Similarity 99.8; Pred. No. 0;
Matches 2639; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3124 cttaagcgattcaicggccgtcgatttacaacgtcgactgggaaacccggcgtt 3183
Db 2736 CCTAAATCCAATTCAGTGGCGCTGCTTTTACACGTCGTGACTGGGAAACCCCTGGCGCTT 2677
Qy 3184 acccaacttaatcgcttgacgacatcccccttttcgcagctggcgtaatagcgaaag 3243
Db 2676 ACCCAACTTATTCGCTTGCAGCACATCCCCCTTTCGCCACGCTGGCGTAATAGCGAAGAG 2617

Qy 3244 gccgcgaccgcatgcccttcccaacagttgcgcagcctgaatggcgaatggcgctgatg 3303
Db 2616 GCCCGCACCGATCGCCCTCCCAACAGTTGGCAGCCCTGAATGGCAATGGCCCTGATG 2557
Qy 3304 cggattttctcttaacgcatctgtgagttattcaacccgcataatggtagcactcagt 3363
Db 2556 CGGTATTTCCTTACGCATCTGTCGGTATTTCACACCCGCATATGTGCACCTCTCAGT 2497
Qy 3364 acaatctctctgatgccgcatagttaagccagcccgacaccccgcccaacacccgctgac 3423
Db 2496 ACAATCTGCTGATGCCGCATAGTTAAGCCAGCCCGCACACCCGCCAACACCCGCTGAC 2437
Qy 3424 gcgcctgacgggctgtctgtcccccgcataccgcttacagacaagctgtgacccgtctcc 3483
Db 2436 CGGCCCTGACGGGCTTGTCTGCTCCCGCATCCCGTTACAGACAACAGCTGTGACCCGCTCC 2377
Qy 3484 gggagctgcatgtgcagaggttttcacggtcatccacgaaacgcgcgagagcaagggc 3543
Db 2376 GGGAGCTGCATGTGTGAGAGGTTTTCACCGTCATCACCAGAACGCCGAGAGCAAGGGC 2317
Qy 3544 ctgctgacgcctatttttataggttaatgtcatgataataatgttcttctagacgtca 3603
Db 2316 CTCGTGATACGCCATATTTTATAGGTTAATGTCATGATAAATAAGTTCTTAGACGTCA 2257
Qy 3604 ggtggcacttttcggggaaaatgtgcgcggaacccctatttcttattttcttaatacat 3663
Db 2256 GGTGGCACTTTTCGGGAAATGTGCGCGGAACCCCTATTGTGTTATTTTCTTAATACAT 2197
Qy 3664 tcaaatatgcatcgctcatgagacaataacccctgataaaatgcttcaataatattgaaaa 3723
Db 2196 TCAAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAA 2137
Qy 3724 aggaagatgatgattcaacattccgtgtgcgccttattcccttttttgcgcattt 3783
Db 2136 AGGAAGATGATGATATTTCACATTTCCGTGTCGCCCTATTTCCTTTTTCGGGCATTT 2077
Qy 3784 tgccttctgtttttgtctacccgaaaaacgctgggaaagtaaaagatgctgaaagatcag 3843
Db 2076 TGCCTTCTGTGTTTTCCTCACCCAGAACCGTGGTGAAGTAAAGATGCTGAAGATCAG 2017
Qy 3844 ttgggtgcagagtggtttacatgaaactgcatcacaacgagcggtaagatcccttgagagt 3903
Db 2016 TTGGGTGCACGAGTGGGTTTACATCGAATGATCTCAACACGCGTAAAGATCCTTCAGAGT 1957
Qy 3904 ttgcgcccgaagaacgcttttccaatgatgagcacttttaaagttctgctatgtggcgcg 3963
Db 1956 TTTCGCCCGGGAAGAACGTTTCCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCG 1897
Qy 3964 gtattatccccgtattgacgcccgggcaagagcaactcgcgcgcacatacactattctcag 4023
Db 1896 GTATTATCCCGTATTGACCGCGGCAAGAGCAACTCGGTCGCGGCATACACTATTCTCAG 1837
Qy 4024 aatgacttggtgactactaccagtcacagaagaagcttctacgattggcatgacagta 4083
Db 1836 AATGACTTGGTGTGACTACTACCACTACAGAAAGAGATCTTACGGATGGCATGACAGTA 1777
Qy 4084 agagaattatgctgctgcccataaccatgagtgataacacactgcgcgcaacttactctg 4143
Db 1776 AGAGAATTATGCAGTGTGCTGCCATACCATGAGTATACACTGCGGCCCACTTACTTCTG 1717
Qy 4144 acaacgactcgagagaccgaagagcctaaccgcttttttgcacaacatggggatcatgta 4203
Db 1716 ACAACGATCGGAGGACCGAAGAGCTAAACCGCTTTTTCACAAACATGGGGATCATGTA 1657
Qy 4204 actgccttgatcttggaaccggagcgtgaatgaagccatacccaaacgagagcgtagac 4263
Db 1656 ACTGCCTTGATCGTGTGGGAACCGGAGCTGAATGAAGCCATACCAACAGACGAGCGTGAC 1597
Qy 4264 acccagatgctgtagcaatggcaacaacgcttgccgaactattaaactggcggaactatt 4323
Db 1596 ACCAGATGCCCTGTAGCAATGGCAACACGTTGCCAANACTATTAACTTGGCGAAGTACTT 1537

QY 4324 actctagcttccggcgaacaaataatagactggtatgaggcggaataaaagttgagaccga 4383
Db 1536 ACTTACGTTCCTCCGGCAACAAATTAATAGACTGTGATGAGCGGATAAAGTTGACGACCA 1477
QY 4384 ctctcgcctcggcccttccggtcgtggtttatgctgataaaactgagccgctgag 4443
Db 1476 CTTCCTGGCTCGGCCCTTCGGCGTGTGGTGTATGCTGATAAATCTGGAGCCGGTGAG 1417
QY 4444 cgtgggtctcgggtatcatctcagcactggggccagatggtaagccctccgtatcgtta 4503
Db 1416 CGTGGGTCTCGCGGTATCATCTGAGCACATGGGGCCAGATGGTAAGCCCTCCCGTATCGTA 1357
QY 4504 gttatatacagcagcggggagtcaggaactatggaagaacgaataagacagatcgctgag 4563
Db 1356 GTTATCTACACGACGGGGATCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAG 1297
QY 4564 ataggctcctcactgattgaagcatctgtaactgcagacaaagtcttactcatatatactt 4623
Db 1296 ATAGGTGCCCTCACTGATTAAGCAATTTGGTAAGTCTGACCAAGTTTACTCATATATACTT 1237
QY 4624 tagattgatttaaaacttcatttttaatttaaaggatctagtggaagatccctttttgat 4683
Db 1236 TAGATTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTAT 1177
QY 4684 aatctcatgacaaataatcccttaactgaggttttcttccactgagcgtcagaccccgta 4743
Db 1176 AATCTCATGACCAAAATCCCTTAACGTGATTTTTCCTCCACTGAGCGTCAGACCCCGTA 1117
QY 4744 gaaaagataaaggatctctctgagatcccttttctcgcgtaaatctgctcttgcaa 4803
Db 1116 GAAAGATCAAGAGATCTCTGAGATCCTTTTTTTCTCGCGCTAATCTGCTGCTTGCAA 1057
QY 4804 acaaaaaaccacgcgtaccagcgggtggttttgcggatcaagagctacccaactctt 4863
Db 1056 ACAAAAAACACACCGCTACAGCGGTGTTTGTTCGGGATCAAGAGCTACCAACTCTT 997
QY 4864 ttctcgaaggttaactggttcagcagcagcagacataccaaataactgctcttagtgtag 4923
Db 996 TTTCCGAAGGTAACTGGCTTCAGCAGCGCAGATACCAAAATCTGTTCTCTAGTGTAG 937
QY 4924 cctagttaggccaccactcaagaactctgtagcaccgctacatacctcgtctgcta 4983
Db 936 CCGTAGTTAGGCCACCACTTCAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTA 877
QY 4984 atcctgttaccagtggtcgtcgcagtgagataagtcgtgtcttaccgggttggactca 5043
Db 876 ATCTGTTACCACTGCTGCTGCCAGTGGCGATAAGTCGTCTTACC GGTTGGACTCA 817
QY 5044 agacgatagttaccgataaaggcagcagcgtcggtgagcgggggttcgtgcacacag 5103
Db 816 AGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGTGAACGGGGGTTCGTGCACACAG 757
QY 5104 cccagcttgagcgaacgactacacccagactagatgatacagcgtgagctatgagaa 5163
Db 756 CCAGCTTGGAGGACAGCACTACACCGAACTGAGATACCTACAGCGCTAGCTATGAGAA 697
QY 5164 agcgcacgcttcccgaaagggaagcggacaggtatcccgtaagcggcgggttcgga 5223
Db 696 AGCGCCACGCTTCCCGAAGGGAAGAGCGGACAGGTATCCGGTAAGCGGAGGGTCGGA 637
QY 5224 acagagagcagcagcagcagcgttccagggggaacgctggtatcttattagtcctgctc 5283
Db 636 ACAGGAGCGCAGGAGGAGCTTCAGGGGGAAGACCGCTGTATCTTTATAGTCTCTGC 577
QY 5284 ggggttcgcacactctgactgagcgtcgtatcttctgctcgtcagggggggcggagc 5343
Db 576 GGGTTTCGCCACCTCTGACTTGAGCGTCCATTTTCTGTATGCTCTCGTCAGGGGGGCGGAC 517
QY 5344 ctatgaaaaaacccagcagcagcgttcttctacgggttctcgtgcttttctgctggcctttt 5403
Db 516 CTATGAAAAAACGCCAGCAACGCGGCTTTTACGGTTCCTGCGCTTTTCTGCTGGCCTTT 457
QY 5404 gctcacatgtttcttctcgtggtatccccctgattctcgtggaataacgctattaccgcttt 5463

Db 456 GCTCACATGTTCTTCTCGTTTATCCCTGATTCTCTGGATAACCGTATTACCGCTTT 397
QY 5464 gagtgcctgataccctcgcctcgcagccgaacacgagcagcagtcagtgagcag 5523
Db 396 GAGTGCAGTGTATACCGCTCGCCGACGCCGAAACGACGCGCAGGAGTCACTGACCGAG 337
QY 5524 gaagcgaagagcgcgaataacgaacacgcctctcccgcgcttggcgcttcaataa 5583
Db 336 GAAGCGGAAGAGCGCCCAATATACGAACGCGCTCTCCCGCGCTTGGCGGATTCAATTAA 277
QY 5584 tgcagctgcacagacagcaggtttcccgactggaagcgggagtgagcgaacgaataat 5643
Db 276 TGCAGCTGCGCACACAGGTTTCCGACTGGAAAGCGGCGAGTGAAGCAACGAATTAAT 217
QY 5644 gtgagttagctcactcattagcagcccgcttaccctttacactttatgcttccgctcgtatg 5703
Db 216 GTGAGTTAGCTCACTCATTTAGGACCCCGAGCTTTACATTTTATGCTTCGCGCTCGTATG 157
QY 5704 ttgtgtgaattgtgagcggataacaattttcacacaggaacagctatgacctgattac 5763
Db 156 TTGTGTGGAATTGTGAGCGGATAACAATTTTCACAGGAACAGCTATGACCATGATTAC 97
QY 5764 gcca 5767
Db 96 GCCA 93

RESULT 10
CVU47294/c
LOCUS
DEFINITION Cloning vector pAdvantage<TM>, complete sequence.
ACCESSION U47294
VERSION U47294.2 GI:6854618
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4392)
AUTHORS Groskreutz,D.J. and Schenborn,E.T.
TITLE Increased Gene Expression in Mammalian cell lines-using pAdvantage(TM) DNA as a Co-transfectant
JOURNAL Promega Notes 48, 8-12 (1994)
REFERENCE 2 (bases 1 to 4392)
AUTHORS Groskreutz,D.J. and Schenborn,E.T.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1996) D.J. Groskreutz, R&D, Promega Corporation, 5445 East Cheryl Parkway, Madison, WI 53711, USA
REFERENCE 3 (bases 1 to 4392)
AUTHORS Grooms,K.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Technical Writing, Promega Corporation, 2800 Woodshollow Road, Madison, WI 53711-5399, USA
REMARK Sequence update by submitter
COMMENT On Feb 2, 2000 this sequence version replaced gi:1200457.
FEATURES
source
1..4392
Location/Qualifiers
misc_feature
1..1725
/note="adenovirus DNA insert"
misc_RNA
780..939
/product="adenovirus VAI RNA"
misc_RNA
1036..1198
/product="adenovirus VAI RNA"
gene
complement(2904..3764)
/gene="Ampr"
CDS
complement(2904..3764)
/note="ampicillin resistance"
/codon_start=1
/product="beta-lactamase"
/protein_id="AAA89081.1"

/db_xref="GI:1200458"
translation="MSQHRVALIPFAAFCLPVPFHPBTLVKVKAEDQLGARVGY
IEDLNSGKILESFRPERPMMSTFKVLICGAVLSRIDAGQQLGRRIHYSQNDLVE
YSPVTEKHLFDGHTVRELCSAAITMSDNTAANLLITIGGPKELTAPLHNMGDHVTRL
DRPELNEAIPIIDERDTTPVAMATTLRLKLTGELITLASRQOLIDMWEADKVGPL
LRSLPAGWFIADKSGAGERSGLIIAALPGDKPSRIVIYITGSAQTDWERNRQIA
EIGASLIKHW"
BASE COUNT 973 a 1177 c 1265 g 977 t
ORIGIN

Query Match 45.7% Score 2635.4; DB 12; Length 4392;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2636; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3131 cgaattcactgcccgtcgcttttacaacgtcgtagctgggaaacccctggcgttaccacaac 3190
DB 4366 CGAATTCACTGGCGCTGCTTTTACAACGTCGTGACTGGGAACCCCTGGCGTTACCCNAAC 4307
QY 3191 ttaatcgcccttgagcagacatccccctttccagctggcgtaataagagagagcccgca 3250
DB 4306 TTAATCGCCTTGCAGCACATCCCCCTTTGCGCAGCTGGCGTAAATAGCGAGAGGCCGCA 4247
QY 3251 ccgatcccttcccaacagttgcagcctgaatggcgaatggcgcctgatcggtatt 3310
DB 4246 CCGATCGCCCTTCCCAACAGTTGCGAGCCTGAATGGCGAATGGCGCTGATGCGGTATT 4187
QY 3311 ttctccttaecgcatctgtgcggtatttcaaccgcaatggtgcacttcagtaacaatct 3370
DB 4186 TTCTCCTTACGCATCTGTGCGGTATTTACACACCGCATATGTCGACCTCTCAGTACAATCT 4127
QY 3371 gctctgatgcgcatagtttaagcagccccgcagacccgcgcgaacccgcgtgacgcgcct 3430
DB 4126 GCTCTGATGCGCGCATAGTTAAAGCAGACCCCGACACCCCGCCAACCCCGCTGACGCGCCT 4067
QY 3431 gacgggctgtctgtctccggcatccgttacagacaagctgtgacgcgtctccggagct 3490
DB 4066 GACGGGCTGTCTGCTCCCGCATCCGCTTACACACAGCTGTACCCGTCTCCGGAGCT 4007
QY 3491 gcatgtctagagaggttttaccgctcatccagaaacgcgcgagacgaagggccctctga 3550
DB 4006 GCATGTGTACAGAGGTTTTCACCGTCATCACCGAAACCGCGAGACGAAGGCGCTCGTGA 3947
QY 3551 tacgcctattttataggttaagtcatgataataagttttcttagacgtcaggtggca 3610
DB 3946 TACGCCATATTTTATAGGTTAATGTATGATTAATAGTTTCTTAGACGTCAGGTGGCA 3887
QY 3611 ctttccggggaatgtgcgcggaacccctattgtttatttttcaaatatcatcaata 3670
DB 3886 CTTTTCCGGGGAATGTGCGCGGAACCCCTATTTGTTTATTTTTTCTAAATACATTTCAATA 3827
QY 3671 tgtatccgctcatgagacaataaaccttgataaaatgcttcaataataattgaaaaagggaaga 3730
DB 3826 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA 3767
QY 3731 gtatgagtattcaacatttcogtgcgccttattcccttttttgcggcatttgccttc 3790
DB 3766 GTATGAGATTTCACATTTCCGCTGCGCCCTATTTCCTTTTTCGGGCATTTTGCCTTC 3707
QY 3791 ctgtttttgcctcccccagaaacgcgtgtgaaagttaaagatgctgaagatcagttggtg 3850
DB 3706 CTGTTTTTGTCTACCCAGAAACCGTGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG 3647
QY 3851 cacagtggttataatcgaaactgattcacaacgcggttaagatccttgagagttttcgcc 3910
DB 3646 CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTCGAGAGTTTCGCC 3587
QY 3911 ccgaagaacgctttcccaatgatgagcacttttaagttctctatgtgcgcggtattat 3970
DB 3586 CCGAAGAACGTTTTTCCAAATGATGAGCACTTTTAAAGTTCCTGATGTGGCGCGGTATTAT 3527
QY 3971 cccgtattgcgcgggcaagagcaactcggctcgcccatcacatatttctcagaatgact 4030

DB 3526 CCCGTATTGACCGCGCGCAAGAGCAACTCGGTCGCCCGCATACACTATTCTCAGAAATGACT 3467
QY 4031 tggttgagtactaccagtcacagaaaaagcatcttacgagatggcatgacagtaagagaat 4090
DB 3466 TGGTTGAGTACTACCAGTCAACAGAAAGCATCTTACGSGATGGCATGACAGTAAGAGAT 3407
QY 4091 tatcgagtctgcgcataaaccatgagtataacactgcggccaacttactcttgacaacga 4150
DB 3406 TATGCAGTCTGCCATACCATACCATGAGTATAACACTCGCGCCCAACTTACTTCTGACACGA 3347
QY 4151 tcgagagaccgaagagcctaaccgcttttttgcacaacatgggggatacatgtaactcgc 4210
DB 3346 TCGGAGGACCGAAGAGCTAAACCGCTTTTTCACAACATGGGGGATCATGTAACTCGCC 3287
QY 4211 ttgatcgttgggaacccggagctgaatgaagcatalaccaaaacgagcgtgacaccaga 4270
DB 3286 TTGATCGTTGGNAACCGGAGCTGAATGAAGCATACCAAAACGACGAGCGGTGACACCAGA 3227
QY 4271 tgcctgtagcaatggcaacaacgcttgogcaaacatttaacttgggaactattactctag 4330
DB 3226 TGCCTGTAGCAATGCAACAACGTTGCGCAAACTATTAACTGGGGAACACTACTTACTCTAG 3167
QY 4331 cttcccggaacaattaatagactggatggagcgggataaaagtgcgagaccacttctgc 4390
DB 3166 CTTCGCGCAACAATTAATAGACTGGATGGAGGCGGATAAAAGTTGACAGGACCACTTTCG 3107
QY 4391 gctcgcccttcccgctggctggtttattgctgataaaatctggagccggtgagcgtgagt 4450
DB 3106 GCTCGGCCCTTCCGCTGGCTGGCTGTTATTGCTGATAACTCTGGAGCCGCTGAGCGTGGGT 3047
QY 4451 ctcgcggtatcattgcagcactggggccagatggtgaagccctcccgatctcgtagtattct 4510
DB 3046 CTCGCGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 2987
QY 4511 acagcaggggagtcagcgaactatgataacaaaatagacagatcctcagatagatg 4570
DB 2986 ACACAGCGGGAGTCAAGCAACTATGATGAACAAATAGACAGATCCTGAGATAGGTG 2927
QY 4571 cctcactgattaaagcattggtaactgtcagaccgaagtttactcatatatacttagattg 4630
DB 2926 CTTCACTGATTAAAGCATTTGGTAACGTGCAGACCAAGTTTACTCATATATATCTTTAGATTG 2867
QY 4631 atttaaaacttcatttttaatttaaaagatcctagggtgaagatccttttgaataatctca 4690
DB 2866 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATATCTCA 2807
QY 4691 tgacaaaatcccttaacgtgagtttctccactaagcgtcagacccctagaaaaaga 4750
DB 2806 TGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGCTCAGACCCCGTAGAAAAA 2747
QY 4751 tcaagagatctcttgagatcctttttctgcgcgtaactctgctgcttgcaaacaaaaa 4810
DB 2746 TCAAGGATCTTCTTGAGATCCTTTTTTCTGCGGTAATCTGCTGTGCAAAACAAAA 2687
QY 4811 aacacgcgtcacccggcggtgtgttgcgggatcaagagctaccaactctttttccga 4870
DB 2686 AACCAACGCTACCAAGCGGTGTTGTTTGCGGGATCAAGAGCTACCAACTCTTTTTCGA 2627
QY 4871 aggttaactgcttcagcagacgcagataccaaatcttcttcttctagctagtcgtagt 4930
DB 2626 AGGTAACTTGGGTTTTCAGCAGACGCGAGATACCAATACTTGTTCTTCTAGTGTAGCGGTAGT 2567
QY 4931 taggccaccacttcaagaactctgtagcacgcgcctacatactcgtcgtgtaacctgt 4990
DB 2566 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTCTGT 2507
QY 4991 taccagtggtcgtgcacagtgccagtaagtgtcttctaccgggttggtgactcaagaagat 5050
DB 2506 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT 2447
QY 5051 agttaccggaataagggcagcagcgtcggtgaacggggggttcgtgacacagccagct 5110
DB 2446 AGTTACCGGATAAGGCGCAGCGGCTGAACGGGGGGTTCGTGTGCACACAGCCCCAGCT 2387

Db 2182 TGACGGGCTGCTGCTCCCGGCATCCGCTTACAGACAMGCTGTGACCGTCTCCGGGAGC 2241
 Qy 3490 tgcattgtcagaggtttttccaccgtcatcaccgaaacgagcagaaagggcctcgtg 3549
 Db 2242 TGCATGTGTACAGAGTTCACCGCTCATCACCAGAAACGCGAGACGAAAGGGCTCGTG 2301
 Qy 3550 atagcctatttttataagggttaattgcatgataataatggtttcttagacatcagggtggc 3609
 Db 2302 ATACGCCATATTTTATAGGTTAAATGTGATGATAATAATGGTTTCTTAGACATCAGGTGGC 2361
 Qy 3610 acttttgggggaaatgtgcgaggaacccctatttcttatttttctaaatacatcctaatt 3669
 Db 2362 ACTTTTGGGGGAATGTGCGCGGAACCCCTATTGTTTATTTTCTTAATACATCAAT 2421
 Qy 3670 atgtatccgtcatgagacaaataaccctgataaaatgcttcaataatattgaaaagggaag 3729
 Db 2422 ATGTATCCGCTCATGAGACAATAACCCGTGATAATGCTTCAATAATATTGAAAAGGAAG 2481
 Qy 3730 agtatgatttcaacatttccgttgtcgccttattcccttttttttgaggcaattttgctt 3789
 Db 2482 AGTATGAGTATTCAACATTTCCGTTGTGCGCCTTATTCCTTTTGTGGCGCATTTTGGCCTT 2541
 Qy 3790 cctgttttctcaccagaaacgctggtgaaagtataaaagatgctgaagatcagttgggt 3849
 Db 2542 CCTGTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGT 2601
 Qy 3850 gcaagtggtttacatcgaactggatctcacaacgagcgtgaagatccttgaggttttgcg 3909
 Db 2602 GCACGAGTGGGTTACATCTGAACACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCG 2661
 Qy 3910 ccggaagaacgttttccaatgatagacactttttaaagtctgctatgttggcgaggtatta 3969
 Db 2662 CCCGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCGTATGTGGCGCGGTATTA 2721
 Qy 3970 tccggtattgacgcccgaagagcaactcgttcgcgcgcatacacactattctcagaatgac 4029
 Db 2722 TCCGCTATTGACGCGGGCAAGACACTCGTTCGCCCATACACTATTCACAGAAATGAC 2781
 Qy 4030 ttggttgagtactaccagtcacagaaagcatcttaccggtgcatgacagagtagagaa 4089
 Db 2782 TTGGTTGAGTACTCACAGTACAGAAAGCATCTTACGGATGGCATGACAGTAAAGAGAA 2841
 Qy 4090 ttatgcagtgctgccataaccatgagtgataaacactgcgcccaacttacttctgacaacg 4149
 Db 2842 TTATGCAGTGTGCGCATAAACATGAGTATAACACTGCGGCCAACTTACTTCTGACAAAG 2901
 Qy 4150 atcggaggaacgaaggagctaacgcgttttttgcacaacatgggggatacatgttaactcgc 4209
 Db 2902 ATCGGAGACCGAAGGAGCTAACCAGCTTTTGGACAACTGGGGGATCATGTAACTCGC 2961
 Qy 4210 cttgatcgttgggaacccggagctgaatgaagccataccaaacgacgagcgtgacacacag 4269
 Db 2962 CTTGATCTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCAAG 3021
 Qy 4270 atgcctgtagcaatggcaaacacttgccgcaaaactattaaactggcggaactacttactcta 4329
 Db 3022 ATGCGTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGGAATACTTACTCTA 3081
 Qy 4330 gcttcccggaacaaataataagatggtgagggcggaataaagttgcaaggaccacttctg 4389
 Db 3082 CTTCCCGGCAACAAATTAATAGATGATGATGAGGCGGGAATAAGTTGACAGACCACTTCG 3141
 Qy 4390 cgcctggcccttccggctggctgtttattgtctgataaaatcgtgagccggtgagcgtggg 4449
 Db 3142 CGCTCGGCCCTTCCGGCTGGCTGTTTATTTGCTGATTAATCTGGAGCCGGTGAGCGTGGG 3201
 Qy 4450 tctcgcggtattcattgagcaactggggccagatggttaagccctcccgctatcgtagttatc 4509
 Db 3202 TCTCGCGGTATCATTTGACGCACTTGGGGCCAGATGGTAAAGCCCTCCCGTATCGTAGTTATC 3261
 Qy 4510 tacacgaggggagtcaggcaactatggatgaacgaataagacagatcgcgtgagataggt 4569
 Db 3262 TACACGAGGGGGGTACAGCAACTATGGATGAACGAATAAGACAGATGCCTGAGATAGGT 3321

Qy 4570 gctcactgattaaagcattggttaactgacagaccaagtttactcatatatacttttagatt 4629
 Db 3322 GCTTCACTGATTAAGCATTTGCTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATT 3381
 Qy 4630 gatttaaaacttcatttttaatttaaaaggatctaggtagaagatccttttttgataatctc 4689
 Db 3382 GATTAAACATTCATTTTAAATTTAAAGAGATCTAGGTGAAGATCCTTTTGTGATAATCTC 3441
 Qy 4690 atgacccaaaatcccttaacgtgagtttctcactgagcgtcagaccccgtagaagaag 4749
 Db 3442 ATGACCAAAATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTACAGCCCGTAGAAAG 3501
 Qy 4750 atcaaaagatctcttgagatcccttttttctgagcgttaactctgctgcttgcacaaacaaa 4809
 Db 3502 ATCAAAAGATCTTCTTTGAGATCCTTTTCTGCGCGTAACTGCTGCTTGCACAAACAAA 3561
 Qy 4810 aaacaccgcgttacagacggttggtttgttgcggatacaagagctaccaactcttttccg 4869
 Db 3562 AAACACCGCTTACCAGCGGTGTTTGTGTTGCGGATCAAGAGCTACCAACTCTTTTTCGG 3621
 Qy 4870 aaggttaactggtctcagcagagcagacatacacaatactctcctctagtgtagcgtag 4929
 Db 3622 AAGTTAACTGCTTTCAGCAGAGCGCAGATACCAATACTGCTTCTTAGTGAGCCGTAG 3681
 Qy 4930 ttaggccaccacttcaagaactctgtagcacgcctacatacctctgctctcttaactcgt 4989
 Db 3682 TTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTCTGCTAAATCTCTG 3741
 Qy 4990 ttaccagtggctgtgcccagtgccagtgccagtaagtctgtcttaccgggttggaactcaagacga 5049
 Db 3742 TTACCACTGGCTGTGTCAGTGGCGATTAAGTCTGTCTTACCGGTTGGACTCAAGACGA 3801
 Qy 5050 tagttaccgggataaggcagcagcgtcgggtgaaacgggggttctgtgcacacagccagc 5109
 Db 3802 TAGTTACCGGATTAAGCCAGCGTCCGGCTGAACGGGGGTTCTGTGCACACAGCCCAAGC 3861
 Qy 5110 ttgagcgaaacgaactacacgaactgagatcacctacagcgtgagctatgagaaagcgc 5169
 Db 3862 TTGGAGCGAACGACCTACACCGAACTGAGATACCTTACAGCTGAGCTATGAGAAAGCGCC 3921
 Qy 5170 acgcttccggaaggagaaagcggagacaggtatccggtaacgagcagggctcggaacagga 5229
 Db 3922 ACGTTCCCGAAGGAGAGAACGGACAGGTATCCGGTAAAGCGGAGGTTGGAAACAGGA 3981
 Qy 5230 gagcgacgagggagcttccaggggaaacgcctgttattcttagtccctgtcggttt 5289
 Db 3982 GAGCGCAGGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGTCTCTGTCTGGGTTT 4041
 Qy 5290 cgccacctctgacttgagcgtcgatttttgtgctgctcaggggagggcgagcctatgg 5349
 Db 4042 CGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCGTCAAGGGGCGGAGCCCTATGG 4101
 Qy 5350 aaaaacgcagcaaacagcggccttttaccggttctggtccttttctggtccttttctcac 5409
 Db 4102 AAAAACGCGCAACAGCGCGCTTTTACGGTTCCTGGCCTTTTGTGGCCTTTTCTCTCAC 4161
 Qy 5410 atgttcttctcgtgtataccctgattctgtgataacccgtattaccgcttttagtga 5469
 Db 4162 ATGTTCTTCTCGGTTATCCCTTGATTTCTGTGATTAACCGTATTATCCGCCCTTTGAGTGA 4221
 Qy 5470 gctgataccgctcgcgagcgaacaccgagcagcagcagcagcagcagcagcagcagcagc 5529
 Db 4222 GCTGATACCGCTCGCGCAGCGAAGCACCAGCGCAGCGAGTCACTGAGCGAGGAAGCG 4281
 Qy 5530 gaagagcgcacaaacgcacacccctctcccgcgcttggccgattcattaaatgcagc 5589
 Db 4282 GAAGAGCGCCCAATACGCAAAACCGCTCTCCCGCGCGCTTGGCCGATTCATTATATGACG 4341
 Qy 5590 tggcacacagaggttccgactgaaacggcgagtgagcgaacgaatcaatgtagt 5649
 Db 4342 TGGCACGACGAGGTTTCCCGACTTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGT 4401

```
QY 5650 taqctcaactcattagcaccacccaggcttttacacttttatctcttcaggctcgatgtttgat 5709
Db 4402 TAGCTCACCTCATTAGCACCACCCAGGCTTTTACACTTTTATGCTTCCGGCTCGTATGTTGTG 4461
QY 5710 ggaattgtgaggggataaacaatttcacagagaacagctatgacctattacgcc 5766
Db 4462 GGAATTGTGAGCGGATAACAAATTTCACAGAGGAACAGCTATGACCATGATTACGCC 4518

RESULT 12
LOCUS U02456 4519 bp DNA circular SYN 29-MAR-1996
DEFINITION Cloning vector pRAJ275, complete sequence.
ACCESSION U02456
VERSION U02456.1 GI:413822
KEYWORDS Cloning vector pRAJ275.
SOURCE Cloning vector pRAJ275.
ORGANISM Cloning vector pRAJ275.
REFERENCE 1 (bases 1 to 4519);
AUTHORS artificial sequence; vectors.
TITLE Cloning vector pRAJ275.
JOURNAL Cloning vector pRAJ275.
AUTHORS Kitts, P.A.
TITLE ClONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 4519)
Jefferson, R.A.
TITLE Plant reporter genes: the GUS gene fusion system
JOURNAL (in) Setlow, J.K. (Ed.);
GENETIC ENGINEERING: PRINCIPLES AND METHODS;
10: 247-263 (1988)
REFERENCE 3 (bases 1 to 4519)
AUTHORS Kitts, P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. This vector is no longer
available from CLONTECH and CLONTECH will not update or revise this
sequence.

FEATURES
source 1..4519
/dbxref="taxon:31825"
BASE COUNT 1130 a 1131 c 1190 g 1068 t
ORIGIN

Query Match 45.7%; Score 2635.4; DB 12; Length 4519;
Best Local Similarity 100.08; Pred. No. 0;
Matches 2636; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3130 gccaattcactggcgcgtcttttacaacgtctgactgggaaacccctggcgttacccaa 3189
Db 1883 GCGAATTACTGGCGCTGCTTTTACAACGCTGCTGACTGGAAACCCCTGGCGTTACCCAA 1942
QY 3190 cttaatgccttgacgacacatcccccttcgcagctggcgtaatagagaagcccgcc 3249
Db 1943 CTTAATGCGCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGC 2002
QY 3250 accgatgcctctcccaacagttgcagcctggaatggcgaatgcgcctgatcggtat 3309
Db 2003 ACCGATGCGCCCTCCCAACAGTGTGCACGCTTGAATGGGGAATGGCGCTGATGCGGTAT 2062
QY 3310 ttctctctacgcatactgtgcggtatttcacacccgcataatggtgcactctcagtaaatc 3369
Db 2063 TTTCTCTTACGATCTGTGCGGTATTTTACACCGCATATGTTGCACCTCTCAGTACAATC 2122
QY 3370 tgcctgtgcgcgcatagttaagcagccccgcacaccccgcaaccccgctgacgcgcc 3429
Db 2123 TGCTCTGATGCGCGCATAGTAAAGCCAGAGCCCGCCGACACCCCGCCGCTGACGCGCC 2182
QY 3430 tgacgggctgtctgtctcccgcatcctgtacagacaagctgtgacgctctccgggagc 3489
```

```
Db 2183 TGACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGC 2242
QY 3490 tgcattgtcagaggttttcaocgctcatccacgaacacgcgcagacgaagaagggcctcgtg 3549
Db 2243 TGCAATGTGTCAGAGGTTTTCACCGTCAATCAGCAAAACGCGAGAGCAAAAGGCCCTCGTG 2302
QY 3550 atacgcctatttttataggttaagtcaatgataaataatggtttcttagacgcctcaggtggc 3609
Db 2303 ATACGCCCTATTTTATAGGTTAATGTCATGAATAATGTTCTTTAGACGTCAGGTGGC 2362
QY 3610 actttccggggaaatgtgcgcgaacccctattgtttattttcttaataatatacattcaaat 3669
Db 2363 ACTTTTCGGGGAAATGTGCGGGAACCCCTATTTGTTATTTTCTTAATATACATTCAAAT 2422
QY 3670 atgtatccgcctcatgagacaataaacccctgataaatgcttcaataatattgaaaaaggaag 3729
Db 2423 ATGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTTGAAAAAGGAAG 2482
QY 3730 agtatgagattcaacatttccgctgcgccttattcccttttttttgcggcattttgcctt 3789
Db 2483 AGTATGAGTATTCAACATTTCCGCTGTGCGCCCTTATTTCCCTTTTTCGCGCATTTTGCCTT 2542
QY 3790 cctgtttttgtccaccagaaacgcgtgtgaaagttaaagataaagatgctgaagatcagttgggt 3849
Db 2543 CTTGTTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGT 2602
QY 3850 gcacgagtgggtttacatcgaaactggatctcaacagcggtaagatccttgagagttttcgc 3909
Db 2603. GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGC 2662
QY 3910 ccggaagaacgcttttccaatgatgagcacttttaaaagtctctatgtggcgcggtatca 3969
Db 2663 CCGGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCGCGGTATTTA 2722
QY 3970 tccgtattgacgcgcgggaacagcaactcgtcgcgcacatacactattctcagaatgac 4029
Db 2723 TCCGCTATTGACGCGCGGGAAGAGCAACTCGGTGCGCGCATACACTATTTCTCAGAAATGAC 2782
QY 4030 ttggttgagtactcaccagctcacagaaaagcatcttcacggatggcagacagtaagaaa 4089
Db 2783 TTGGTTGAGTACTACACAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA 2842
QY 4090 tiatgcagtgctgccataaaccatgagtataacactgcggcccaacttacttctgacaacg 4149
Db 2843 TTATGCACTGCTGCCATAACCATGAGTGATAACACTTACGCGGCAACTTACTTCTGACAACG 2902
QY 4150 atcggaggccgaagagcgttaacccgttttttgcaacaactgggggacatgttaactcgc 4209
Db 2903 ATCGGAGGACCGAAGGAGCTAACCCGCTTTTTCACACACATGGGGGATCATGTAACTCGC 2962
QY 4210 cttgatcgttgggaacccggagcgtgaatgaagccaataccaaacgacgagcgtgacaccag 4269
Db 2963 CTTGATCGTTGGGAACCGGAGCTGAATGAAGCAATACCAACAGCAGCAGCGTGACACCAG 3022
QY 4270 atgctgtagaataggcaaacacgcttgccaaactattaaactggcgaaactacttactcta 4329
Db 3023 ATGCTGTAGCAATGGCAACAACGTTTGCACAACTATTAACTTGGCAACTACTTACTCTTA 3082
QY 4330 gcttcccgccgaacaattaatagactggatggaggggataaaagtgtgcaggaaccactctg 4389
Db 3083 GCTTCCCGCAACAAATTAATAGACTGGATGGAGCGGGATTAAGTTGCAGAGCAACCTCTG 3142
QY 4390 cgtcggcgcccttccgctggctgggtttattgtctgataaaactctggagccggtgagcgtgg 4449
Db 3143 CGCTCGGCCCTTCCGGCTGGCTGGTTTATTTGCTGATAAATCTGGAGCCGCTGAGCGTGGG 3202
QY 4450 tctcgggttatcatgtcagcaactggggcgagatgtaagccctcccgctatcgtatgtatc 4509
Db 3203 TCTCGCGGTATCATGTGACACTGGGGCGAGATGGTAAGCCCTCCCGTATCTGATGTTATC 3262
QY 4510 tacagacggggagtcaggcaactatggatgaacgaataagacagactcgtctgagatagct 4569
Db 3263 TACACGCGGGGATCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGT 3322
```

Qy	4570	g	c	t	c	a	c	t	g	a	a	c	a	t	t	g	t	a	a	c	t	g	t	c	a	a	c	a	a	g	t	t	a	c	t	a	b	a	t	a	c	t	t	a	g	a	t			
Db	3323	G	O	C	T	C	A	C	T	G	A	T	T	A	A	G	C	A	T	T	G	T	A	C	T	G	T	A	A	G	C	A	T	T	A	A	G	C	A	T	T	A	A	G	C	A	T			
Qy	4630	g	a	t	t	a	a	a	a	c	t	t	a	a	t	t	a	a	a	a	a	g	a	t	c	t	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a		
Db	3383	G	A	T	T	A	A	A	A	C	T	C	A	T	T	T	A	T	T	A	A	A	G	A	T	C	T	A	A	A	G	A	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
Qy	4690	a	t	a	c	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a			
Db	3443	A	T	G	A	C	C	A	A	A	T	C	C	T	T	A	A	C	G	T	G	A	T	T	T	C	G	T	T	C	C	A	T	T	T	C	C	A	T	T	T	C	C	A	T	T	T	T		
Qy	4750	a	t	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a			
Db	3503	A	T	C	A	A	A	G	A	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
Qy	4810	a	a	a	c	a	c	e	c	q	t	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	a	c	
Db	3563	A	A	C	C	A	C	C	E	T	A	C	A	C	C	G	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
Qy	4870	a	a	g	t	a	a	c	t	y	g	g	t	t	c	a	g	a	g	c	a	g	a	t	c	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a			
Db	3623	A	A	G	T	A	A	C	T	G	G	T	T	C	A	G	A	G	C	G	A	T	A	C	A	A	A	T	A	C	T	G	T	T	C	T	T	A	G	T	G	T	T	C	T	A	G	T		
Qy	4930	t	a	g	g	c	c	a	c	a	c	t	t	a	a	a	c	t	t	a	c	a	c	c	c	t	a	c	a	c	a	c	t	a	c	t	c	a	c	t	c	a	c	t	c	a	c	t	c	
Db	3683	T	T	A	G	G	C	C	A	C	T	T	C	A	A	G	A	C	T	C	T	A	G	A	C	C	G	C	T	A	C	T	A	C	T	T	G	C	T	T	G	C	T	T	A	T	T	C	T	
Qy	4990	t	t	a	c	a	g	t	g	g	t	g	c	t	g	g	e	a	t	a	a	g	t	c	t	t	a	c	c	g	g	t	t	g	a	c	a	a	c	a	a	c	a	a	c	a	a	c	a	a
Db	3743	T	T	A	C	C	A	G	T	G	S	C	T	G	C	C	A	T	A	A	G	T	C	G	T	C	T	T	A	C	C	G	G	T	T	G	G	A	C	T	C	A	A	C	A	C	A	C	A	

Qy	5650	tagctcaactcattaggacccccaggctttacacctttatgcttcggctcgtatgttgt 5709
Db	4403	TAGCTCACTCAATAGGCACCCAGGCTTTACACTTTATGCTTCGGCTCGTATGTTGTT 4462
Qy	5710	ggaattgagcgcgatacaaatattcacacaggaacagcctatgaccatgattacgcc 5766
Db	4463	GGAAATTGTAGCGGATACAATTTACACAGGAACAGCTATGACCATGATTACGCC 4519
RESULT 13		
XXU02437		
LOCUS		
Cloning vector pT3/T7-LUC, 4674 bp DNA circular SYN 29-JAN-1997		
DEFINITION		
U02437		
ACCESSION		
U02437.1 GI:413803		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
Cloning vector pT3/T7-LUC.		
Cloning vector pT3/T7-LUC.		
artificial sequence: vectors.		
1 (bases 1 to 4674)		
Kitts,P.A.		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424- 8222 or (800) 662-2566, extension 3. This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH; this vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.		
FEATURES		
source		
1..4674		
/organism="Cloning vector pT3/T7-LUC"		
/db_xref="taxon:31848"		
BASE COUNT 1239 a 1092 c 1141 g 1202 t		
ORIGIN		
Query Match 45.7%; Score 2635; DB 12; Length 4674;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 2635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	3133	aattcaactggccgctgttttacaacgtcgtgactgggaaaaacctggcgttacccaactt 3192
Db	2008	AATTCACTTGGCCGCTGCTTTTCAACGTCGTGACTGGGAAAACCCCTGGCGTTTACCAACTT 2067
Qy	3193	aatgccttgagacacatcccccttccagctggcgtaataagcgaagagccgcacc 3252
Db	2068	AATGCCCTTGAGCAATCCGCCCTTCCGACGTGGCGTAAATACGGAAGAGCCCGCACC 2127
Qy	3253	gatgccttcccaacagtgcgacgcctgaatggcgcaatggcgctgatcggtatttt 3312
Db	2128	GATGCCCTTCCAAACAGTTGCGAGCCTGAATGGCGAATGGCGCTGATCGGTATTTT 2187
Qy	3313	ctccttagcctatctgtgggtattttcaacacgcgatattggtcaactctcagtacaattctgc 3372
Db	2188	CTCCTTAGCGCATCTGTGCGGTATTTTACACCGCATATGGTGACTCTCAGTACAACTGTC 2247
Qy	3373	ctctgtagcgcatagtttaagcagccccgcacacccgcacaccccgctgacgcgccttga 3432
Db	2248	TCTGATGCGGCATAGTTAAGCAGCCCCGACACCCGCCCAACACCGGTGACGCGCCCTGA 2307
Qy	3433	cgggcttctgctccccggcatccgcttacagacaagctgtgacgctctccggagctgc 3492

|||||
Db 2308 CGGGCTGTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 2367
Qy 3493 atggttcagaggttttcacgcgtcatcacccgaaacgcgcgagacgaagggcctcggtgata 3552
|||||
Db 2368 ATGTGTACAGAGGTTTTCACCGTCATCACCGAAACGCCGAGACGAAGGGCCTCGTGATA 2427
Qy 3553 cgcctattttataggttaagtcatgataataatgggtttctttagacgtcaggtggcact 3612
Db 2428 CGCCTATTTTATAGGTTAATGTCATGATAATAATGTTTCTTAGACGTCAGGTGGCACT 2487
Qy 3613 ttctgggggaaatgtgcgggaaacccctattgtttatttttctaaatacatccaatatg 3672
Db 2488 TTTTCGGGGAATATGTCGGCGAAACCCCTATTGTGTTATTTTCTAAATACATTCAAAATATG 2547
Qy 3673 tatcgcgtcatgagacaataaaccttgataaatgcttcaataatattgaaaaaggaagagt 3732
Db 2548 TATTCGGCTCATGAGACAATTAACCGTGATAATGCTTCAATAATATTTGAAAAAGGAAGAGT 2607
Qy 3733 atgagttatcaacatttcgcgtgcgccttattcccttttttggcgcaattttgccttcc 3792
Db 2608 ATGAGTATTCAACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGGCATTTTGCCTTCCT 2667
Qy 3793 gtttttgcacccagaaacgcgtggtgaaagtaaaagatgctggaagatcagttgggtgca 3852
Db 2668 GTTTTTCCTTCACCCAGAAACCGTGTGTAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 2727
Qy 3853 cgagtgggttacatcgaaactggatctcaacacgcggtaagatccttgagagttttccccc 3912
Db 2728 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC 2787
Qy 3913 gaagaacgttttccaatgatgagcaacttttaagttctgctatgtgycggtattatcc 3972
Db 2788 GAAGAAGCTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTTATTATCC 2847
Qy 3973 cgtattgcgcgggaaagagaacactcgtgcgcgcatacactatttctcagaatgacttg 4032
Db 2848 CGTATTGACCGCGGGCAAGAGCAACTTCGGTTCGGCGCATACATTTCCTCAGAATGACTTG 2907
Qy 4033 gttgagtactcacagatcacagaaaagcatcttacggatggcatgacagtgaagagaatta 4092
Db 2908 GTTGTAGTACTACCAGTCACAGAAAAGCATCTTACGGATGGCATGACATGAAGAGAATTA 2967
Qy 4093 tgcagtgtgcataaccatgagtgatgaatacactgcggccaacttactcttgacaacgatac 4152
Db 2968 TGCAGTGTCTGCCATACCATAGATGATATAACACTGCGGCCAACTTACTTCTGACAACGATC 3027
Qy 4153 ggaagccgaagagatcaaccgttttttgcacaacatgggggatacatgtaactgcctt 4212
Db 3028 GGAGGACCGAAGAGGCTAACCCGCTTTTGTGACAACTGGGGGATCATGTAACTCGCCCTT 3087
Qy 4213 gatcgttgggaaacggagctgaatgaagccataccaaacgaagcgtgacacccagatg 4272
Db 3088 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTCACACCAGATG 3147
Qy 4273 cctgtagaactggcaaacactgtgcgaactatttaactggcgaaactattactctagct 4332
Db 3148 CCGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTTGGCGAACTACTTACTCTAGCT 3207
Qy 4333 tcccgccaacatttaatactgagtgagcgagataaagttagcaggaacacttctgcgc 4392
Db 3208 TCCCGCAACAATTAATAGCTGGATGGAGGCGGATAAAGTTGCGAGGACCACTTCGCGC 3267
Qy 4393 tcggcccttcggctggctggtttattgctgataaactctggagccgggtgagcgtgggtct 4452
Db 3268 TCGGCCCTTCCGGCTGGCTGGTTATTGTGATAAATCTGGAGCCGCTGAGCGTGGGTCT 3327
Qy 4453 cgcggtatcatgtcagcaactggggcgagatggtgaagccctcccgatctagttatctaac 4512
Db 3328 CCGCGGTATCATGTGACACTGGGGCCAGATGGTTAAGCCCTCCCGTATCGTAGTTATCTAC 3387
Qy 4513 acgacgggagtcaggcaactatgtagaacgaataagacagatcgcctgagatagtgcc 4572
|||||

Db 3388 ACACGGGAGTCAGCAACTATGTGATGAACGAATAAGACAGATGCTCTGAGATAGGTGCC 3447
Qy 4573 tcaactgataagcattggttaaactgtcagaccaagtttactcatatatactttagattgat 4632
Db 3448 TCACTCGATTAAAGCATTTGGTAACGTGTGACACCAAGTTTACTCATATATACTTTAGATTGAT 3507
Qy 4633 ttaaaacttcaatttttaatttaaaaggatctagggtgaagatccttttttgataatctcatg 4692
Db 3508 TTAATAACTTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATG 3567
Qy 4693 accaaatccctttaacgcgtgagtttttcgttccactgagcgtcagaccccgtagaaaaagatc 4752
Db 3568 ACCAAATCCCTTAAACGTGAGTTTTCGTTCCACTGAGCGTCAAGCCCGTAGAAAAAGATC 3627
Qy 4753 aaaggatctcttgagatcctttttctgcgcgtaactgtgcgtctgcaacaacaaaaa 4812
Db 3628 AAAGGATCTTCTTGAGATCCTTTTTCGCGGTAACTGCTGCTTGCAACAAAAANA 3687
Qy 4813 ccaccgctaccagcgggtgttbtgtttgcgggatcaagagctaccacactctttttccgaag 4872
Db 3688 CCACCGCTACCAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCGGAAG 3747
Qy 4873 gtaactgggttcagcagagcagataccaaatctcttctctctagtgtagcgtagtta 4932
Db 3748 GTAACGTGGCTTCAGCAGAGCGCAGATACCAAAATACGTCTCTTAGTGTAGCCGTAGTTA 3807
Qy 4933 ggccacacttcaagaactctgtagcacgcctacatacctgcgtctgtaactctgtta 4992
Db 3808 GGCCACCACCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTCTTGA 3867
Qy 4993 caagtgcgtgcgcgtggcgataagtgctcttaaccgggttggaactcaagacgatatg 5052
Db 3868 CCAGTGGCTGTGCTCAGTGGCGATAAGTGTGTCTTACCGGGTTGGACTCAAGACGATAG 3927
Qy 5053 ttaccggataaaggcagcgcgtcgagctaaacgggggttcgtgcacacagcccagcttg 5112
Db 3928 TTACCCGATTAAGCGCGCAGCGTGGCTGAAACGGGGGTTCTGTGCACACAGCCAGCTTG 3987
Qy 5113 gaggcaacgacctacaccgaactgagatacctacagcgtgagctatgaaagcgcacg 5172
Db 3988 GAGCGAAGCAGCTTACACCGAATGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAGC 4047
Qy 5173 cttcccgagggaagaaaggcgcagaggtatccggtaagcgggaaggctcggaacagagag 5232
Db 4048 CTTCCGAGAGGGAAAGCGGACAGGTATCCGCTAAGCGCAGGTCGGAGCAGGAGAG 4107
Qy 5233 cgcacgagagagcttccaggggaaacgcctggtactctttagtccctgtcggtttcgc 5292
Db 4108 CGCACAGGAGGAGCTTCCAGGGGAAACGCTGCTATCTTTATAGTCTCTGTGGGTTTCGC 4167
Qy 5293 cactctgactgagcgtcgtatttttctgtagctcgttcaggggggggcggagcctatgaaa 5352
Db 4168 CACCTCTGACTTGAGCGTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCTATGNNAA 4227
Qy 5353 aacgcgcaaacgcgccttttttaacggttctcgtgccttttgcgttccttttgcctcatg 5412
Db 4228 AACGCAGCAACACGCGCCTTTTACGGTTCTCTGGCCCTTTTGTGGCTTTTGTCTCATG 4287
Qy 5413 ttcttctcgtcttccccctgattctgtggataaacgtaataccgccttttgagtgagct 5472
Db 4288 TTTCTTTCTTGGCTTATCCCTGATTCTGTGGATTAACCGTATTAACCGCTTTTGTAGTGAGCT 4347
Qy 5473 gataccgctcgcgcgacgaacgacgcagcagcagtcagtgagcaggaagcgga 5532
Db 4348 GATACCGCTCGCGCAGCGCAACGACCGCAGCAGCAGGTGAGTGAGCAGGAGGAGCGMA 4407
Qy 5533 gagcgcccaatacgaacacgcctctcccgcgcttcccgagcttccatttaagtcagctgg 5592
Db 4408 GAGCGCCCAATAGCGAAACCGCCTCTCCCGCGCGTGGCCGATTTCATTAATCAGCTGG 4467
Qy 5593 cagcagaggttccccgactggaaagcggcagtgagcgaacgaatttaagtgagttag 5652
Db 4468 CACGACAGGTTTCCCACACTGGAAGCGGCGAGCTGACGGCAACCAATTAATGTAGTTAG 4527

Qy 5653 ctcaactattaggacccccaggcttttacacttttatgtctccggctcgtatgtgtgga 5712
|||||
Db 4528 CTCACCTATTAGGACCCAGCGCTTACACTTTATGCTTCGCGCTCGTATGTTGTGGA 4587
|||||
Qy 5713 attgtgagcgataacaatttcacacaggaacagctatgaccatgattacgccca 5767
|||||
Db 4588 ATTTGAGCGGATACAAATTTACACAGGAACAGCTATGACCATGATTACGCCA 4642
|||||

RESULT 14
AF276982
LOCUS AF276982 5027 bp DNA linear SYN 06-JUL-2000
DEFINITION Integrative vector pNLE1, complete sequence.
ACCESSION AF276982
VERSION AF276982.1 GI:8927566
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Salvatore, P., Cantalupo, G., Pagliarulo, C., Tredici, M., Lavitola, A.,
Bucci, C., Bruni, C.B. and Alifano, P.
TITLE A new vector for insertion of any DNA fragment into the chromosome
of transformable *Neisseriae*
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5027)
AUTHORS Salvatore, P., Cantalupo, G., Pagliarulo, C., Tredici, M., Lavitola, A.,
Bucci, C., Bruni, C.B. and Alifano, P.
TITLE Direct Submision
JOURNAL Submitted (12-JUN-2000) Biology, University of Lecce, Via
Monteroni, Lecce, LE 73100, Italy
FEATURES
source
1. 5027
Location/Qualifiers
/organism="Integrative vector pNLE1"
/db_xref="taxon:130490"
/lab_host="Neisseria meningitidis"
/note="Integrative vector for transformable *Neisseriae*"
<32..733
/note="contains partial leucyl-tRNA synthetase (leus)
coding region"
762..806
/note="neisserial uptake sequence and putative
transcription terminator"
970..>1513
/gene="dam"
970..>1513
/gene="dam"
/note="DNA methylase"
/codon_start=1
/transl_table=11
/product="DNA-adenine-methylase"
/protein_id="AAF82121.1"
/db_xref="GI:8927567"
/translation="MASKPFLKAGGKKLVPEIHNLPPTKRLLEPCGSAALS
ALPEHYLLINDADLGLRLKEKSGFIDYARSFISENDSRFEYLRQFNFS
KNHRSALYILNHRFNGLCRYNSGVNVPGRYKSPYFPQOEMEGETQKSDRVE
LMCGDFQMLSDTNDTDTVC"
1605..2339
/gene="ermC"
1605..2339
/gene="ermC"
/note="RNA methylase"
/codon_start=1
/transl_table=11
/product="erythromycin resistance"
/protein_id="AAF82122.1"
/db_xref="GI:8927568"
/translation="MNEKNKHSQNFITSKHINDKIMTNIRLNHDNIFEIGSGKH
TLELVQCNFVTAIEIDHKLCKTTEKNLVHDNFQVLNKDILQKFPKNQSYKIFGNI
PYNSTDIIRKIVPDSIADEIYLIVEYGFARLNLTKRSLALMAEVDISLSMVR
EYHPKRVNSLSLRNKKRSIRISHKDKQYNYFVMKWNKYEKKIIFTKQFNLSLKH
AGIDDLNISEFQFLSLFNSYKLFNK"

BASE COUNT 1424 a 1203 c 1127 g 1273 t
ORIGIN
Query Match 45.7%; Score 2634.4; DB 12; Length 5027;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3131 cgaattcactggcgcgtctgttttacaacgctcgtgactgggaaacccctggcgttaccacaac 3190
|||||
Db 2392 CGAATTCACCTGGCGCTCGTTCATCAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAAC 2451
|||||
Qy 3191 ttaatcgcttgcagcacatcccccttccgacagctgcgtaataagcagaagagcccgca 3250
|||||
Db 2452 TTAATCGCCTTTCAGCACATCCCCCTTTCCGACGTGGCGTAATAGCGAAGAGGCCGCA 2511
|||||
Qy 3251 ccgatcgcccttcccaacagttgcgacgctcgtgaaaggcgaatggcgctgagcggtatt 3310
|||||
Db 2512 CCGATCGCCCTTCCCAACAGTTGCGCAGCTCAATGGCGAATGGCGCTGATCGCGTATT 2571
|||||
Qy 3311 ttctcttacgcatctgtgcggtatttcacacccgcataatggcgactctcagtaacaatct 3370
|||||
Db 2572 TTCTCCTTACGCATCTGTGCGGTATTTCCACACCGCATATGGTGCACCTCTCAGTACAATCT 2631
|||||
Qy 3371 gctctgatgcgcgatagttaagcagcccgacaccccgcaaccccgctgacgagccct 3430
|||||
Db 2632 GCTCTGATGCCGCAATAGTTAAAGCCAGCCCGCACACCCCGCTGACGCGCCCT 2691
|||||
Qy 3431 gacggcgttctgtctcccgccatccgcttacagacagctgtagccgtctccgagagct 3490
|||||
Db 2692 GACGGCCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 2751
|||||
Qy 3491 gcatgtgcagaggttttcacccgtcatcccgaaacgcgcgagacgaagggccctcgtga 3550
|||||
Db 2752 GCATGTGCAGAGGTTTTCACCGTCATCACGGAACGCGGAGAGAGGAAGGSCCTCGTGA 2811
|||||
Qy 3551 tacgcctattttataggttaatgcatgataataatggttctcttagcgcgtcagtgcca 3610
|||||
Db 2812 TACGCGTATTTTATAGGTTAATGTCATGATAAATGGTTTCTTTAGACGTCAAGTGCCA 2871
|||||
Qy 3611 ctttcggggaaatgcgcggaacccctattgttttcttaataacatacaataca 3670
|||||
Db 2872 CTTTTCGGGGAAATGTGCGGGAACCCCTATTTGTTTATTTTCTTAATACATTCAAATA 2931
|||||
Qy 3671 tgtatccgctcatgagacaataaacctgataaaatgcttcaataatattgaaaagaaga 3730
|||||
Db 2932 TGTATCCGCTCATGAGACAATAAACCTGTATAATGCTTCAATAATATTGAAAAGAGAGA 2991
|||||
Qy 3731 gtagagtattcaacatttcggtgctgcgccttattcccttttttgcgcatttgccttc 3790
|||||
Db 2992 GTATGAGTATTCAACATTTCCGCTGTCGCCCTTATTTCCCTTTTTCGCGCATTTTCCCTTC 3051
|||||
Qy 3791 ctgttttgcctcaccagaaacgcgtggtgaagtaaaagatgctgaagatcagttgggtg 3850
|||||
Db 3052 CTGTTTTTGTCTACCCACCAACCGCTGGTGAAAGTAAAGATGCTGAAGATCACTTTGGGTG 3111
|||||
Qy 3851 cacgagtggtttacatcgaaatggaatcacaacagcggttaagatccttgagagcttttcgcc 3910
|||||
Db 3112 CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTGAAGATCCTTGAGAGCTTTTCGCC 3171
|||||
Qy 3911 ccgaagaacggttttccaatgatgagcaacttttaaaagtctgctgtagcgaggtattat 3970
|||||
Db 3172 CCGAAGAAGCGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGATTTAT 3231
|||||
Qy 3971 cccgtattgacgcgggcaagagcaactcgttcgcgcatacacatattctcagaatgact 4030
|||||
Db 3232 CCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACATTTCTCAAGATGACT 3291
|||||
Qy 4031 tggttgagtactcaccagtcacagaaaaagcatcttcacggatggcatgacagtaagaat 4090
|||||
Db 3292 TGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAT 3351
|||||
Qy 4091 tatgcagtgctgcataaccatagtgatataacactgcggccaaacttaactcttgacaaaga 4150
|||||

[illegible]

D	b	4432	CGCTTCCCGAAGGAGAAAGCGGCACAGGTATCCGGTAAGCGCAGGGTCGGAACAAGGAG	4491
Q	y	5231	agcgacagaggagcttccagggggaacgcctggatcttttatagtcctgtcgggttcc	5290
D	b	4492	AGCGCACGAGGAGCTTCCAGGGGGAACGCTGGTATCTTTATAGTCTCGGTTC	4551
Q	y	5291	gccaccttgacttagcgctgatatttttgtatgctcgtcagggggcgagcctatgga	5350
D	b	4552	GCCACCTTGACTTAGCGCTCGATTTTTTGTATGCTCGTCAGGGGGCGGAGCCTATGGA	4611
Q	y	5351	aaaagcgcagcaacgcgccctttttaagcgttctcctggcccttttggccttttgcctcacaa	5410
D	b	4612	AAAAAGCCAGCAACGCCGCCCTTTTACGGTTCCCTGGCCTTTTGTGGCCTTTTCTCACA	4671
Q	y	5411	tgtcttctcgcgttatcccctgatctctggaataaccgtattaccgcctttgagtgag	5470
D	b	4672	TGTTCTTCTTCGCGTTATCCCTCATTTCTGTGGATAACCGTATTACGCCCTTTGAGTGAG	4731
Q	y	5471	ctgataccgctcgccgcagcgcgaacccgagcgcagcagtcagtgcagcgagaagcgg	5530
D	b	4732	CTGATACCGCTCGCGCAGCGCAACGACCGAGCGCAGCGAGTCAGTGAGCGAGAAGCGG	4791
Q	y	5531	aagagcgccaataacgcaaacccgcctctccccgcgcgcttggccgatcatctaagcagct	5590
D	b	4792	AAGAGCGCCCAATAACGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATSCAGCT	4851
Q	y	5591	ggcacagaggttcccgactcgaaagcgggcagtcagcacaagcgaattaatgatgt	5650
D	b	4852	GGCAGCACAGGTTTCCCAGACTGGAAAGCGGCATGTAGCGCACGCAANTTAATGTAGTT	4911
Q	y	5651	agctcaactcattagagaccgccgctttacactttatgcttcgcgcctcgatatgtgtg	5710
D	b	4912	AGCTCACTCATTAGGCACC GCCAGGCTTTACACTTTATGCTTC CGGCTCGTATGTTGTG	4971
Q	y	5711	gaattgtgacgggataacaatttacacaggaacagcctatgaccatgattacgcc	5766
D	b	4972	GAATTGTGACGGGATAACAATTTCCACAGGAACAGCTATGACCATGATTACGCC	5027
RESULT 15				
BACKNB/c				
LOCUS		5647 bp DNA linear BCT 14-APR-2000		
DEFINITION		Bacillus stearothermophilus pSTK1 and PUBI10 genes for ORFs and kanamycin nucleotidyltransferase, complete cds.		
VERSION		D29979		
KEYWORDS		ORF; kanamycin nucleotidyltransferase; kanamycin resistance; thermostable mutation.		
SOURCE ORGANISM		Bacillus stearothermophilus DNA. Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Geobacillus. 1 (sites) Liao,H., McKenzie,T. and Hageman,R. Isolation of a thermostable enzyme variant by cloning and selection in a thermophile Proceedings of the National Academy of Sciences of the United States of America. 83 (3), 576-580 (1986)		
REFERENCE		86120983		
AUTHORS		MEDLINE PUBLISHED 3003740		
TITLE		2 (sites) Narumi,I., Nakayama,N., Nakamoto,S. and Kihara,H. Complete nucleotide sequence of pSTK1, a cryptic plasmid from Bacillus stearothermophilus TK015 Biotechnol. Lett. 15, 1013-1016 (1993)		
JOURNAL		3 (sites) Narumi,I., Nakayama,N., Nakamoto,S., Kimura,T., Yanagakisawa,T. and Kihara,H. Construction of a new shuttle vector pSTE33 and its stabilites in Bacillus stearothermophilus, Bacillus subtilis, and Escherichia coli Biotechnol. Lett. 15, 815-820 (1993)		
REFERENCE		4 (sites)		
JOURNAL				
AUTHORS				
TITLE				

AUTHORS	Narumi,I., Nakayama,N., Nakamoto,S. and Kihara,H.
TITLE	Bacillus stearothermophilus plasmid pstK1 replicon is functional in Escherichia coli
JOURNAL	Biotechnol. Lett. 17, 475-480 (1995)
REFERENCE	5 (bases 1 to 5647)
AUTHORS	Nakamoto,S.
JOURNAL	Unpublished (1995)
FEATURES	Location/Qualifiers
source	1..5647
	/organism="Geobacillus stearothermophilus"
	/db_xref="taxon:1422"
CDS	243..518
	/note="pstK1"
	/citation-[2]
	/codon_start=1
	/transl_table=11
	/evidence=not_experimental
	/product="ORF1"
	/protein_id="BAA06246.1"
	/db_xref="GI:849038"
	/translation="MKKIGIRITEENDKKLREIATRYGMSINSVASFIIGKWLDENYG LKDRILEGLAASQKEVMERVAQVVEELFSPNEKIKELERHGLKLNF"
rep_origin	524..713
	/note="pala minus"
	/citation-[2]
	/evidence=not_experimental
CDS	538..768
	/note="pstK1"
	/citation-[2]
	/codon_start=1
	/transl_table=11
	/evidence=not_experimental
	/product="ORF2"
	/protein_id="BAA06247.1"
	/db_xref="GI:849038"
	/translation="MPRKLARVRVLGTSRKALPCICVQAFSRQRLPAPADKKHAS WGPVPRGNNAERPVIGGIRIKGKNSSETPVV"
CDS	1013..1822
	/note="pstK1"
	/citation-[2]
	/codon_start=1
	/transl_table=11
	/evidence=not_experimental
	/product="ORF3"
	/protein_id="BAA06248.1"
	/db_xref="GI:849040"
	/translation="MGIHVMTGCGCRFELTSTINWYELVLYVEYEVNITRLDVA VDDEKGYFINTLYKKLKDDEVTSPFKKHAIENIVIEGTEIGHTLYFGAPSSDIQV RYENKVGMDIDVNNRTETIQLDRDRAHVVAQIITADDPLGEIVAGLLRNYIQRT CLNTEKORFIDVLAEGASKLTAKDLQVINOFSKKNITYDEMIKIIRQSK"
CDS	2168..2929
	/note="kanamycin resistance gene
	pub110"
	/codon_start=1
	/transl_table=11
	/evidence=experimental
	/product="kanamycin nucleotidyltransferase"
	/protein_id="BAA06249.1"
	/db_xref="GI:484047"
	/translation="MNGPITMTREERMKIVHEIKERILDKYGGDVKAIGVYGLGRQT DPGSDIEMKCYMSTEARFSEHWTGKWKVNFYSEIILDIYASQVSDWPLTHGQ FSYILPIYDSGGYLEKRYTAKSVEAQKPHDAICLIVELFEYAGKWNIRVQGGPT DFPKLTVOVAMAGAMLIGLHHRICYTTSASVLTEAVKQSDLPISGYDHLQCQFVMSGOLS DSEKLLLESLENWNGIQWTERHGVIYDYSKRIPP"
variation	2405
	/citation-[1]
	/phenotype="thermostable mutant"
	/replace="g"
	2556
variation	
	/citation-[1]
	/phenotype="thermostable mutant"
	/replace="c"

|||||
Db 4675 TATGCAAGTGTGCAATTAACCATGAGTGTATTAACACTGCGGCCAACTTACTTCTGACAAACGA 4616
QY 4151 tcggagagaccgaaggagctaacgcgttttttgcacaacatgggggatcatgttaactcgc 4210
Db 4615 TCGGAGGACCGAAGGAGCTAACCGCTTTTGTGCACAACATGGGGGATCATGTAACTCGCC 4556
QY 4211 ttgatcgttgggaacccggagctgaatgaagccataccaaacgacgagcgtgacacacga 4270
Db 4555 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGA 4496
QY 4271 tgcctatagaatggcaaacacttcgcacaactattaaactggcgaacttaactctag 4330
Db 4495 TGCTGTAGCAATGGCAACACCTTGGCAAACTATTAACTGGCGAATCTACTTACTCTAG 4436
QY 4331 ctctccgggcaacaattaaatagactggatggagcggataaagtgcaggagaccacttcgc 4390
Db 4435 CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGACGAGCACTTCTGC 4376
QY 4391 gctcgccctccggctgctggtttattatgtctgataaactgagccggtgagcgtgggt 4450
Db 4375 GCTCGGCCCTTCCGGCTGGCTGTTTATTGCTGATAAATCTGGAGCGGCTGAGCGTGGT 4316
QY 4451 ctgcggttatctgcagcaactgggcccagatggttaagccctccgtatcgttatct 4510
Db 4315 CTGCGGTATCATGTGAGCACTGGGCGCCAGATGGTAAGCCCTCCCGTATCGTGTAGTTATCT 4256
QY 4511 acacgacggggagtcaggcaactatgatagaacgaatagacagatcgctgagatagg 4570
Db 4255 ACAGACGGGGAGTCAGGCAACTATGGATGAACGAANAATACAGAGATCGGTGAGTAGGTC 4196
QY 4571 cctcaatgattgaacttggttaactgacagcaacttaactcataatacttagattg 4630
Db 4195 CCTCACTGATTAGCATTTGGTAATCTGACAGCAAGTTTACTCATATATACTTTAGATTG 4136
QY 4631 atttaaaactctatttaattaaagatctagtggaagatcccttttttgataactca 4690
Db 4135 ATTAAAACTTCAATTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTTGTGATAATCTCA 4076
QY 4691 tgacaaaaatcccttaacgtgagtttctccactgagcgtcagaccccgtagaaaaa 4750
Db 4075 TGACCAAAATCCCTTAACGTGAGTTTCTGTTCCACTGAGCGTCAGACCCGCTAGAAAAGA 4016
QY 4751 tcaaaagatcttttgagatcccttttttctgcgcgttaactgctgtcgtgcgaacaaaa 4810
Db 4015 TCAAAGGATCTTCTGAGATCTCTTTTCTGCGCGTAATCTGCTGTGCAAAACAAAA 3956
QY 4811 aaccacgctacacgagcgtgtgtttgttccggatcaagagctaccaactcttttccga 4870
Db 3955 AACCCCGCTTACCAGCGGTGGTGTGTTGTCGGGATCAAGAGCTTACCACCTCTTTTCCGA 3896
QY 4871 aggttaactggcttcagcagagcagataccaaaactgtcctctctagtgtagccgtagt 4930
Db 3895 AGGTAACCTGGCTACACAGCGCAGATACCAAAATCTGCTTCTAGTGTAGCCGTAGT 3836
QY 4931 tagggcaacacttcaagaactctgtagcccgctacatacactcgtcgtcgttaactcgt 4990
Db 3835 TAGGCCACCACCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAACTCCTGT 3776
QY 4991 taccagtgctgcagtgagcgaatcgtgtcttaccgggttggaactcaagcgtat 5050
Db 3775 TACCAGTGGCTGCTGCCAGTGGCGGATAAGTGTGTCTTACCGGGTGGGACTCAAGACGAT 3716
QY 5051 agttaccggaataaggcgcagcggctgggctgaacgggggttcgtgcacacagcccgct 5110
Db 3715 AGTTACCGGATTAAGGCGCAGCGGTGGGCTGAACGGGGGTTCGTGCACACAGCCAGCT 3656
QY 5111 tggaggaacacactacacccgaactgagatacctacagcgtgagctatgagaaagccca 5170
Db 3655 TGGAGCGAAGCAGCTACACCGAACTGAGATACCTACAGCGTGAAGATTGAGAAAGCGCCA 3596
QY 5171 cgcctccgaaggagaaaggcagagtatccggttaagcggcaggggtcggaacaggag 5230
|||||

Db 3595 CGCTTCCCAGAGGAGAAAGGCGGACAGGTATCCGGTAAGCGCCAGGGTCGGAAACAGGAG 3536
QY 5231 agcgcacggaggagcttccaggggggaaacgcctgttatctttatagtcctgtcgggtttc 5290
Db 3535 AGCGCAGGAGGAGCTTCCAGGGGAAACGCTTGGTATCTTTATATCTCTGTCTGGGTTTC 3476
QY 5291 gccacctcgaacttgagcgtcgatattttgtgactcgtcgtaaggggggcggagcctatgga 5350
Db 3475 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGA 3416
QY 5351 aaaaacccagcaacgcgcttttttacgggttctcctggccttttctggccttttgcacaa 5410
Db 3415 AAAACGCCAGCAACGGGCCCTTTTACGGTTCTCTGGCCTTTTGTGCCCTTTTGTCTACA 3356
QY 5411 tgtcttccctgcgttatcccttgattctgttgataaacgattaccgccttttgagtgag 5470
Db 3355 TGTTCCTTCTCTGGTTATCCCTTGATTCGTGTGATAACCGTATTACCGCCTTTGAGTGAG 3296
QY 5471 ctgataccgctcgcgcagccgaacacgcgagcgcagtcagtcagtcagtcagtcagtcagtc 5530
Db 3295 CTGATACCGCTCGCGCAGCCGAACACGAGCGAGCGAGTCAGTCAGCGAGGAAGCGG 3236
QY 5531 aagagcgcacaatacacgcaaacgcctctcccgcgcttggccgattcattaatcagct 5590
Db 3235 AAGAGCGCCCAATACGCAAAACCCCTCTCCCGCGCGTTGGCCGATTCATTAAATGCACT 3176
QY 5591 ggacgacaggtttcccgactggaaagcgggagtgagcgcaacgcaattaaatgtgagtt 5650
Db 3175 GGCACGACAGGTTTCCCGACTGGAAAGCGGCAAGTCAGCGCAACGCANTTAATGTGAGTT 3116
QY 5651 agtcactcaataggcaaccccggtttacaactttatgcttcgggtcgtatgtgtgtg 5710
Db 3115 AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 3056
QY 5711 gaattgtgagcggaatacaaatttcacaggaacagctatgacatgattacgca 5767
Db 3055 GAATTGTGACGGATAACAATTTTCACACAGGAAACAGCTATGACCATGATTACGCCA 2999
|||||

Search completed: August 31, 2002, 18:10:48

Job time: 16101 sec